# Best Available Copy

(19)

Canadian Intellectual Property Office Office de la Propriété Intellectuelle du Canada

(11) CA 2 303 526

(13) A1

An Agency of industry Canada Un organisme d'Industrie Canada

(40) 03.10.2000 (43) 03.10.2000

(12)

(30)

(21) 2 303 526

(51) Int. Cl.7:

C12N 7/02, C12N 5/10,

C12N 15/51

(22) 31.03.2000

199 15 178.4 DE 03.04.1999

(72)

BARTENSCHLAGER, RALF (DE).

(71)

BARTENSCHLAGER, RALF, Nach dem Alten Schloss 22 D-55239, GAU-ODERNHEIM, XX (DE).

(74)

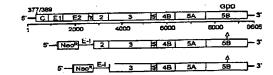
**BORDEN LADNER GERVAIS LLP** 

(54) SYSTEME DE CULTURE DU VIRUS DE L'HEPATITE C

(54) HEPATITIS C VIRUS CELL CULTURE SYSTEM

(57)

The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).



# OPIC OFFICE DE LA PROPRIÉTÉ INTELLECTUELLE DU CANADA

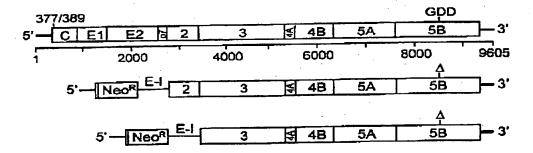
# (12)(19)(CA) Demande-Application

CIPO
CANADIAN INTELLECTUAL
PROPERTY OFFICE

(21) (A1) **2,303,526** 

(43) 2000/10/03

- (72) BARTENSCHLAGER, RALF, DE
- (71) BARTENSCHLAGER, RALF, DE
- (51) Int.Cl.<sup>7</sup> C12N 7/02, C12N 15/51, C12N 5/10
- (30) 1999/04/03 (199 15 178.4) DE
- (54) SYSTEME DE CULTURE DU VIRUS DE L'HEPATITE C
- (54) HEPATITIS C VIRUS CELL CULTURE SYSTEM



(57) The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).

### Hepatitis C Virus Cell Culture System

## **ABSTRACT**

5

10

The hepatitis C virus (HCV) cell culture system according to the invention consists of human hepatoma cells, which are transfected with a HCV-RNA construct, that comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as a minimum of one marker gene for selection (selection gene).

#### Hepatitis C Virus Cell Culture System

#### DESCRIPTION

The invention relates to a hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing transfected HCV specific genetic material, which means they are transfected with HCV specific genetic material.

10

The hepatitis C virus (HCV) is one of the main causes worldwide of chronic and sporadic liver diseases. The history of most HCV infections does not involve any obvious clinical signs, but 80 - 90 % of the infected people become chronic carriers of the virus and 50 % of these chronic carriers of the virus develop chronic hepatitis with 20 % of the different degrees of severity. Approx. chronically infected develop a cirrhosis of the liver over 10 to 20 years, based on what a primary hepatocellular carcinoma can develop. Nowadays chronic hepatitis C is the main indication for liver transplantation. A specific therapy does not exist until now. The only therapy high-dose administration currently available is Interferon alpha or a combination of Interferon alpha and the purine nucleoside analogue Ribavirin. However, only approx. 60 % of all treated persons respond to this therapy and with these, a new viraemia occurs in more than half of all cases after the discontinuation of the treatment.

Due to the high prevalence, especially in industrialized countries, the serious effects of chronic infections and the non-existence of a specific therapy, the development of a HCV specific chemotherapy is an important goal of pharmaceutical research and development. The main problem lies in the previous lack of a suitable cell culture system, which enables the study of virus replication and pathogenesis in eukaryotic cells.

Due to the small amount of virus in blood or tissue, the lack of suitable cell culture systems or animal models (the chimpanzee is still the only possible experimental animal) as well as the lack of efficient systems for producing virus-like particles, it was not possible up to now, to analyze the molecular composition of the HCV particle indepth and to solve it. The information currently available can be summarized as follows: HCV is an enveloped plusstrand RNA virus with a particle diameter of 50 - 60 nm and 10 a medium density of 1.03 - 1.1 g/ml. It was molecularly cloned and characterized for the first time in 1989 (Choo et al., 1989, Science, 244, 359 - 362). The HCV-RNA has a length of approx. 9.6 kb (= 9600 nucleotides), a positive polarity and comprises one open reading frame (ORF), which encodes a linear polyprotein of approx. 3010 amino acids (see Rice 1996, in Virology, B. N. Fields, D. M. Knipe, P. M. Howley, Eds. (Lippincott-Raven, Philadelphia, PA, 1996), vol. 1, pp. 931 - 960; Clarke 1997, J. Gen. Virol. 78, 2397; and Bartenschlager 1997, Intervirology 40, 378 and see Fig. 1 A). During the replication of the virus the polyprotein is cleaved into the mature and functionally active proteins by cellular and viral proteases.

20

Within the polyprotein the proteins are arranged as follows 25 (from the amino- to the carboxy terminus): Core-E1-E2-p7-NS2-NS3-NS4A-NS4B-NS5A-NS5B. The core protein is the main component of the nucleocapsid. The glycoproteins E1 and E2 are transmembrane proteins and the main components of the viral envelope. They probably play an important role during 30 the attachement of the virus to the host cell. These three proteins core, E1, and E2 constitute the viral particle and are therefore called structural proteins. The function of the protein p7 is still not clear. The protein NS2 is probably the catalytic domain of the NS2-3 protease, which 35 is responsible for the processing between the proteins NS2 and NS3. The protein NS3 has two functions, one is a

protease activity in the amino terminal domain, which is essential for the polyprotein processing, and the other a NTPase/helicase function in the carboxy terminal domain, which is probably important during the replication of the viral RNA. The protein NS4A is a co-factor of the NS3 protease. The function of the protein NS4B is unknown.

The open reading frame is flanked on its 5' end by a nontranslated region (NTR) approx. 340 nucleotides in length, which functions as the internal ribosome entry site (IRES), and on its 3' end by a NTR approx. 230 nucleotides in length, which is most likely important for the genome replication. A 3' NTR such as this is the object of patent application PCT/US 96/14033. The structural proteins in the amino terminal quarter of the polyprotein are cleaved by host cell signal peptidase. The non-structural proteins (NS) 2 to (NS) 5B are processed by two viral enzymes, namely the NS2-3 and the NS3/4A protease. The NS3/4A protease is required for all cleavages beyond the carboxy terminus of NS3. The function of NS4B is unknown. NS5A, a highly phosphorylated protein, seems to be responsible for the Interferon resistance of various HCV genotypes (see Enomoto et al. 1995, J. Clin. Invest. 96, 224; Enomoto et al. 1996, N. Engl. J. Med. 334, 77; Gale Jr. et al. 1997, Virology 230, 217; Kaneko et al. 1994, Biochem. Biophys. Res. Commun. 205, 320; Reed et al., 1997, J. Virol. 71, 7187), and NS5B has been identified as the RNA-dependent RNA polymerase.

20

First diagnostic systems have been developed from these findings, which are either based on the detection of HCV specific antibodies in patient serum or the detection of HCV specific RNA using the reverse transcription polymerase chain reaction (RT-PCR), and which are (must be) routinely used with all blood and blood products and/or according to the regulations.

Since the first description of the genome in 1989 several partial and complete sequences of the HCV have been cloned and characterized using the PCR method. A comparison of these sequences shows a high variability of the viral genome in particular in the area of the NS5B gene, which eventually resulted in the classification of 6 genotypes, which are again subdivided into the subtypes a, b, and c.

The genomic variance is not evenly distributed over the genome. The 5'NTR and parts of the 3'NTR are highly conserved, while certain encoded sequences vary a lot, in particular the envelope proteins E1 and E2.

10

The cloned and characterized partial and complete sequences of the HCV genome have also been analyzed with regard to appropriate targets for a prospective antiviral therapy. In the course of this, three viral enzymes have been discovered, which may provide a possible target. These include (1) the NS3/4A protease complex, (2) the NS3 Helicase and (3) the NS5B RNA-dependent RNA polymerase. The 20 NS3/4A protease complex and the NS3 Helicase have already been crystallized and their three-dimensional structure determined (Kim et al., 1996, Cell, 87,343; Yem et al., 1998, Protein Science, 7, 837; Love et al., 1996, Cell, 87, 311; Kim et al., 1998, Structure, 6, 89; Yao et al., 1997, 25 Nature Structural Biology, 4, 463, Cho et al., 1998, J. Biol. Chem., 273, 15045). it has not been successful until now with the NS5B RNA-dependent RNA polymerase.

Even though important targets for the development of a therapy for chronic HCV infection have been defined with these enzymes and even though a worldwide intensive search for suitable inhibitors is ongoing with the aid of rational drug design as well as high throughput screening, the development of a therapy has one major deficiency, namely the lack of cell culture systems or simple animal models, which allow direct, reliable identification of HCV-RNA or

HCV antigens with simple methods which are common in the laboratory. The lack of these cell culture systems is also the main reason that to date the comprehension of HCV replication is still incomplete and mainly hypothetical.

5

Although according to the experts, a close evolutionary relationship exists between HCV and the flavi- and pestiviruses, and self-replicating RNAs have been described for these, which can be used for the replication in different cell lines with a relatively high yield, (see Khromykh et al., 1997, J. Virol. 71, 1497; Behrens et al., 1998, J. Virol. 72, 2364; Moser et al., 1998, J. Virol. 72, 5318), similar experiments with HCV have not been successful to date.

15

30

10

Although it is known from different publications that cell lines or primary cell cultures can be infected with high titre patient serum containing HCV, (Lanford et al. 1994, Virology 202, 606; Shimizu et al. 1993, Proceedings of the National Academy of Sciences, USA, 90, 6037 - 6041; Mizutani et al. 1996, Journal of Virology, 70, 7219 - 7223; M. Ikeda et al. 1998, Virus Res. 56, 157; Fournier et al. J. Gen. Virol. 79, 2376 and bibliographical references quoted in here; Ito et al. 1996, Journal of General Virology, 77, 1043 - 1054), these virus-infected cell lines or cell cultures do not allow the direct detection of HCV-RNA or HCV antigens. The viral RNA in these cells can not be detected in a Northern Blot (a standard method for the quantitative detection of RNA) or the viral protein in Western Blot or а immunoprecipitation. It has only been possible to detect HCV replication with very costly and indirect methods. disadvantageous facts show that obviously the replication in these known virus-infected cell lines or cell cultures is completely insufficient.

Furthermore it is known from the publications of Yoo et al. (1995, Journal of Virology, 69, 32 - 38) and of Dash et al., (1997, American Journal of Pathology, 151, 363 - 373) that hepatoma cell lines can be transfected with synthetic HCV-RNA, which are obtained through in vitro transcription of the cloned HCV genome. In both publications the authors started from the basic idea that the viral HCV genome is a plus-strand RNA functioning directly as mRNA after being transfected into the cell, permitting the synthesis of viral proteins in the course of the translation process, and so new HCV particles are (could be) formed. This viral replication, which means these newly formed HCV viruses and their RNA, have been detected through RT-PCR. However the published results of the RT-PCR carried out indicate, that the HCV replication in the described HCV transfected 15 hepatoma cells is not particularly efficient and is not sufficient to measure the quality, let alone the quantity of the fluctuations in the replication rate after an targeted action with prospective antiviral treatments. Furthermore it is prior art (Yanagi et al., Proc. Natl. 20 USA, 96, 2291-95, 1999), that the highly Sci. conserved 3' NTR is essential for the virus replication. This knowledge strictly contradicts the statements of Yoo et al. and Dash et al., who used for their experiments only HCV genomes with shorter 3' NTRs since they did not know 25 the authentic 3' end of the HCV genome.

The object of the present invention is to provide a HCV cell culture system, where the viral RNA self-replicates in the transfected cells with such a high efficiency that the quality and quantity of the fluctuations in the replication rate can be measured with common methodologies usually found in the laboratory after a targeted action with virus and prospective HCV specific antivirals in particular.

35

30

The solution to this problem is to provide a cell culture system of the prior mentioned type, where the eukaryotic

cells are human cells, in particular hepatoma cells, which are preferably derived from a normal hepatoma cell line, but can also be obtained from an appropriate primary cell culture, and where the transfected HCV specific genetic material is a HCV-RNA construct, which essentially comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR preferably in the order mentioned as well as a minimum of one marker gene for selection (selection gene).

10

30

Here and in the following "NTR" stands for "non-translated region" and is a known and familiar term or abbreviation to the relevant expert.

- 15 Here and in the following the term "HCV-RNA construct" comprises constructs, which include the complete HCV genome, as well as those, which only include a part of it, which means a HCV subgenome.
- 20 A preferred variation of the cell culture system according to the invention, which had proven to be worthwhile in practice, is lodged at the DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German collection of Microorganisms and Cell Cultures) in Braunschweig, Germany under the number DSM ACC2394 (laboratory name HuBl 9-13).

With the cell culture system according to the invention an in vitro system is provided for the first time, where HCV-RNA is self-replicated and expressed intracellularly and in a sufficient amount, so that the quantity of the amounts of HCV- RNA as well as the HCV specific proteins can be determined with conventional and reliably precise biochemical measuring methods. This means an almost authentic cell-based HCV replication system is available for the first time, which is urgently needed for the development and testing of antiviral drugs. This test system provides the possibility of identifying potential

targets for an effective HCV specific therapy and developing and evaluating HCV specific chemotherapeuticals.

The invention is based on the surprising finding that efficient replication of the HCV-RNA only occurs in cells if they have been transfected with an HCV-RNA construct, which comprises at least the 5' and the 3' non-translated regions (NTR) and the non-structural proteins (NS) 3 to 5B and additionally a marker gene for selection (selection gene). The structural genes are obviously without great importance for replication, whereas efficient replication of the HCV-RNA apparently only occurs if the transfected cells are subject to permanent selection pressure, which is imparted by the marker gene for selection (selection gene) linked to the HCV-RNA. Consequently the marker (selection gene) seems on one hand to provoke the selection of those cells, where the HCV-RNA replicates productively, and it seems on the other hand to considerably increase the efficiency of the RNA replication.

20

25

30

35

10

15

An object of the invention is also a cell-free HCV-RNA construct, characterized in that it comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR, preferably in the order mentioned, as well as a marker gene for selection (selection gene).

In the present context the terms 5' NTR and NS3 and NS4A and NS4B and NS5A and NS5B and 3' NTR comprise each nucleotide sequence, which is described in the state of the art as the nucleotide sequence for each functional segment of the HCV genome.

By providing a HCV-RNA construct such as this, a detailed analysis of the HCV replication, pathogenesis and evolution in cell culture is possible for the first time. The HCV specific viral RNA can specifically be created as a complete genome or subgenome in any amount, and it is

possible, to manipulate the RNA construct and consequently to examine and identify the HCV functions on a genetic level.

Because all HCV enzymes identified as a main target for a therapy at the moment, namely the NS3/4A protease, the NS3 helicase and the NS5B polymerase, are included in the HCV-RNA construct according to the invention, it can be used for all relevant analyses.

10

An embodiment of the HCV-RNA construct, which has proven to be worthwhile in practical use, stands out by the fact that it comprises the nucleotide sequence according to the sequence protocol SEQ ID NO:1.

15

20

25

35

Further embodiments with similar good properties for practical use are characterized in that they comprise a nucleotide sequence either according to sequence protocol SEQ ID NO:2 or according to sequence protocol SEQ ID NO:3 or according to sequence protocol SEQ ID NO:4 or according to sequence protocol SEQ ID NO:5 or according to sequence protocol SEQ ID NO:6 or according to sequence protocol SEQ ID NO:8 or according to sequence protocol SEQ ID NO:8 or according to sequence protocol SEQ ID NO:9 or according to sequence protocol SEQ ID NO:10 or according to sequence protocol SEQ ID NO:11.

It is possible to provide the HCV subgenomic construct with a 3' NTR, which has a nucleotide sequence so far unknown in the state of art, a nucleotide sequence, which has been selected from the group of nucleotide sequences (a) to (i) listed in the following:

Т

- (b) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAGTCT
  TTTTTTTTC TTTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT
  GGAGC
- 5 (c) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTAATCTTT
  TTTTTTTTTT TTTTTTTTGA GAGAGAGAGT CTCACTCTGT TGCCCAGACT
  GCAGC
- (d) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAGTC

  10 TTTTTTTTT TCTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCCAG

  ACTGGAGT
  - (e) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAGTCT
    TTTTTTTTTT TCTTTTTTTT TGAGAGAGAG AGTCTCACTC TGTTGCCCAG
    ACTGGAGT

15

20

- (f) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAGTCT
  TTTTTTTTTT TCTTTTTTT TTGAGAGAGA GAGTCTCACT CTGTTGCCCA
  GACTGGAGT
- (g) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAGTCT
  TTTTTTTTTT CTTTTTTTT GAGAGAGAG GTCTCACTCT GTTGCCCAGA
  CTGGAGT
- 25 (h) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTTAAT
  CTTTTTTTT TTTTTCCTTT TTTTGAGAGA GAGAGTCTCA CTCTGTTGCC
  CAGACTGGAG T
- (i) ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTT TTTTTAATC

  30 TTTTTTTTT TTTTCTTTTT TTTTTGAGAG AGAGAGTCTC ACTCTGTTGC

  CCAGACTGGA GT

The marker gene for selection (selection gene) included in the HCV-RNA constructs according to the invention is 35 preferably a resistance gene, in particular an antibiotic resistance gene. This has the advantage that the cells transfected with this construct can easily be selected from the non-transfected cells by adding for example the appropriate antibiotic to the cell culture medium in the case of an antibiotic resistance gene.

In the present context 'antibiotic' means any substance, which impedes the non-transfected host cells or the cells, where the HCV-RNA is not replicating efficiently, continuing to live or grow, especially the cell poison Puromycin, Hygromycin, Zeocin, Bleomycin or Blasticidin.

A preferred marker gene for selection (selection gene) and resistance gene, which has proven to be worthwhile in practice, is the neomycin phosphotransferase gene.

An alternative for the antibiotic resistance genes is for example the thymidine kinase gene, which can be used to carry out a HAT selection.

20

10

The marker gene for selection (selection gene), the preferred resistance gene and the most preferred antibiotic resistance gene is preferably positioned in the HCV-RNA construct after the HCV 5' NTR, which means downstream from the 5' NTR and upstream from the HCV reading frame. However, an insertion in the area of the 3' NTR or another site of the HCV genome or subgenome, for example within the polyprotein, is also conceivable.

In another embodiment of the HCV-RNA construct according to the invention the marker gene for selection (selection gene), in particular an antibiotic resistance gene, is linked to the HCV-RNA or HCV genomic or subgenomic sequence via a ribozyme or a recognition site for a ribozyme.

35

This has the advantage, that after the selection of the cells, in which the HCV-RNA is replicating productively,

the resistance gene in the obtained cell clones can be separated from the HCV subgenomic sequence through a ribozyme-dependent cleavage, namely by activating the inserted ribozyme or in the case of a construct with a recognition site for a ribozyme, by transfecting the the cells (for example through ribozyme into transfection of a ribozyme construct or infection with a which the expression vector, into appropriate ribozyme has been inserted). By this means an authentic HCV genomic construct can be obtained without a resistance gene, which can then form authentic infectious virus particles.

Another preferred embodiment of the HCV-RNA construct according to the invention is characterized in that the construct has at least one integrated reporter gene.

In the following a reporter gene means any gene, whose presence can be easily detected with, in general, simple biochemical or also histochemical methods after being transferred into a target organism, which means a gene, that encodes for a protein, which can be easily and reliably detected and quantified in small amounts with the common measuring methods in the laboratory.

25

10

This variation of the HCV-RNA construct has the advantage that the extent of the replication of this construct can be easily and quickly measured with the methods common in the laboratory using the reporter gene product.

30

The reporter gene is preferably a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) or the SEAP gene (secreted alkaline phosphatase gene). This reporter gene and its products, namely the relevant reporter proteins, can be

detected for example using fluorescence, chemiluminescence, colorimetric measurements or by means of immunological methods (for example ELISA).

A surrogate marker gene can also be considered as a reporter gene. In this context it includes those genes, which encode for cellular proteins, nucleic acids or generally for those functions, which are subject to variation depending on the replication of the virus, and which consequently are either suppressed or activated in the cells, in which the HCV or the HCV-RNA construct multiplies. This means, the suppression or activation of this function is a surrogate marker for the replication of the virus or the replication of the HCV-RNA construct.

15

20

25

30

35

The positions of the reporter genes and the marker gene for selection (selection gene) can be selected in such a way, that a fusion protein made from both genetic products will be expressed. This has the advantage that these two genes can be arranged in such a way in the HCV-RNA construct that their two expressed proteins are fused via a recognition sequence for a protease (for example ubiquitin) or via a self-cleaving peptide (for example the 2A protein of the Picornaviruses) at first and will be separated proteolytically later.

These two positions might as well lie apart from each other in such way, that both genetic products are separately expressed (for example in the order: marker or resistance gene — internal ribosome binding site — reporter gene).

In the case of the reporter gene one embodiment has proven to be particularly worthwhile, where the reporter gene is cloned into the open reading frame of the HCV genome or subgenome in such a way that it will only be transferred to an active form after proteolytic processing. The cell culture system according to the invention can be used for various purposes in each of its embodiments. These comprise:

- 5 The detection of antiviral substances. This can include for example: organic compounds, which interfere directly or indirectly with viral growth (for example inhibitors of the viral proteases, the NS3 helicase, the NS5B RNA-dependent RNA polymerase), antisense oligonucleotides, which will hybridize to any target sequence in the HCV-RNA construct (for example the 5' NTR) and will have an direct or indirect influence on the virus growth for example due to a reduction of translation of the HCV polyprotein or ribozymes, which cleave any HCV-RNA sequence and consequently impair virus replication.
  - The evaluation of any type of antiviral substances in the cell culture. These substances can be detected on the isolated purified enzyme for example with 'rational drug design' or 'high-throughput screening'. Evaluation means mainly the determination of the inhibitory features of the respective substance as well as its mode of action.

20

- The identification of new targets of viral or cellular origin for a HCV specific antiviral therapy. If for example a cellular protein is essential for viral replication, the viral replication can also be influenced by inhibiting this cellular protein. The system according to the invention also enables the detection of these auxiliary factors.
- The determination of drug resistance. It can be assumed that resistance to therapy occurs due to the high mutation rate of the HCV genome. This resistance, which is very important for the clinical approval of a substance, can be detected with the cell culture system

5

10

15

according to the invention. Cell lines, in which the HCV-RNA construct or the HCV genome or subgenome replicates, are incubated with increasing concentrations of the relevant substance and the replication of the viral RNA is either determined by means of an introduced reporter gene or through the qualitative or quantitative detection of the viral nucleic acids or proteins. Resistance is given if no or a reduced inhibition of the replication can be observed with concentration of the active substance. The nucleotide and amino acid replacements responsible for the therapy resistance can be determined by recloning the HCV-RNA (for example by the means of RT-PCR) and sequence analysis. By cloning the relevant replacement/s into the original construct its causality for the resistance to therapy can be proven.

- The production of authentic virus proteins (antigens) for the development and/or evaluation of diagnostics.
   The cell culture system according to the invention also allows the expression of HCV antigens in cell cultures. In principle these antigens can be used as the basis for diagnostic detection methods.
- The production of HCV viruses and virus-like particles, in particular for the development or production of therapeutics and vaccines as well as for diagnostic purposes. Especially cell culture adapted complete HCV genomes, which could be produced by using the cell cultur system according to the invention, are able to replicate in cell culture with high efficiency. These genomes have the complete functions of HCV and in consequence they are able to produce infectious viruses.
- The HCV-RNA construct according to the invention by itself can also be used for various purposes in all its embodiments. This includes first of all:

 The construction of attenuated hepatitis C viruses or HCV-like particles and their production in cell cultures:

5

10

Attenuated HCV or HCV-like particles can be created by accidental or purposefully introduced mutations, such as point mutations, deletions or insertions, which means viruses or virus-like particles with complete ability to replicate, but reduced or missing pathogenicity. These attenuated HCV or HCV-like particles can be used in particular as vaccine.

- The construction of HCV-RNA constructs with integrated foreign genes, used for example as liver cell specific 15 vector in gene therapy. Due to the distinctive liver cell tropism of the HCV and the possibility of replacing parts of the genome by heterologous sequences, HCV-RNA constructs can be produced, where for example the 20 structural proteins can be replaced by a therapeutically effective gene. The HCV-RNA construct obtained in this way is introduced into cells preferably by means of transfection, which express the missing HCV functions, for example the structural proteins, in a constitutive 25 or inducible way. Virus particles, carrying the HCV-RNA construct, can be created by means of this method known to the expert under the term 'transcomplementation'. The particles obtained can preferably be used liver infection of cells. Within these therapeutically effective foreign gene will be expressed 30 and will consequently develop its therapeutic effect.
- The detection of permissive cells, which means cells, in which a productive virus growth occurs. For this purpose either one of the HCV-RNA genomic constructs previously mentioned, which is able to form complete infectious viruses, or one of the HCV subgenomes previously

5

10

15

mentioned, which according to the previously mentioned example will be transfected in a cell line first, which expresses the missing functions in a constitutive or inducible way, is used. In each case virus particles are created, which carry a resistance and/or reporter gene apart from the HCV sequence. In order to detect cells, where the HCV is able to replicate, these cells are infected with viruses generated in this way and subject an antibiotic selection or they are examined depending on HCV-RNA construct by means the determining the presence of the expression of reporter gene. Because an antibiotic resistance reporter gene expression can only be established, when HCV-RNA construct replicates, the cells detected in this way must be permissive. Almost any cell line or primary cell culture can be tested in regard to the permissivity and detected in this way.

The cell culture system according to the invention also 20 permits targeted discovery of HCV-RNA constructs for which there is an increase in the efficiency of replication due to mutations. This occurs either by chance, in the context of HCV-RNA replication, or by targeted introduction into the construct. These mutations, leading to a change in the 25 replication of the HCV-RNA construct, are known to experts adaptive mutations. The invention therefore includes a method for obtaining cell culture adapted mutants of a HCV-RNA construct according to the invention following the above description, in which the mutants have 30 increased replication efficiency compared to the original HCV-RNA construct. It further includes a method for the production of mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with increased replication efficiency compared to the original 35 full-length genome or subgenome or **HCV-RNA** HCV-RNA construct, as well as cell culture adapted mutants of HCV-

RNA constructs, HCV-RNA full-length genomes and HCV subgenomes with increased replication efficiency compared to the original constructs, subgenomes or full-length genomes.

5

10

The method according to the invention for the production of cell culture adapted mutants of a HCV-RNA construct according to the invention, in which the mutants have increased replication efficiency compared to the HCV-RNA construct, is characterised in that a cell culture system according to claim 1, in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to claims 4 to 19, is cultivated on/in the selection medium corresponding to the selection gene, that the cultivated cell clones are collected and that the HCV-RNA construct is isolated from these cell clones.

In an advantageous extension of this production method, the isolated HCV-RNA constructs are passaged at least one more time, that is they are transfected in cells of a cell culture system according to claim 1, the thus obtained cell culture system according to claim 1, in which the transfected HCV specific genetic material is the isolated HCV-RNA construct with a selection gene, is cultivated on/in the selection medium corresponding to the selection gene, the cultivated cell clones are collected and the HCV-RNA constructs are thus isolated.

Using this process variation, the quantity of adaptive mutations and hence the degree of replication efficiency in the relevant HCV-RNA constructs can be increased even further.

The method according to the invention for the production of mutants of a HCV-RNA full-length genome or of a HCV-RNA subgenome or of any HCV-RNA construct with increased replication efficiency compared to the original HCV-RNA

full-length genome or subgenome or HCV-RNA construct, has the following features. Using one of the two production methods presented above, a cell culture adapted mutant of a HCV-RNA construct is produced, isolated from the cells, 5 cloned using state of the art known methods and sequenced. By comparing with the nucleotide and amino acid sequence of the original HCV-RNA construct, the type, number and position of the mutations is determined. These mutations are then introduced into an (isolated) HCV subgenome or full-length genome or any HCV-RNA construct, either by site-directed mutagenesis, or by exchange of DNA fragments containing the relevant mutations.

10

25

30

A test can be carried out to determine or verify which mutations actually are responsible for an alteration of replication efficiency, particularly an increase replication. In this test the corresponding nucleotide and/or amino acid changes are introduced into the original HCV-RNA construct and the modified construct transfected in cell culture. If the introduced mutation actually leads to an increase in replication, then for a HCV-RNA construct with a selectable marker gene, the number of resistant cell clones in the artificially mutated construct should be noticeably higher compared to the untreated construct.

In the case of a construct with a reporter gene, the activity or quantity of the reporter should be noticeably higher for the artificially mutated construct compared to the untreated one.

The cell culture adapted HCV-RNA constructs with high replication efficiency according to the invention are characterized in that, through nucleotide or amino acid exchanges, they are derivable from a HCV-RNA construct according to one of the claims 4 to 19, and that they are

obtainable using one of the two production processes presented above.

These cell culture adapted HCV-RNA constructs can be used to produce any HCV-RNA constructs or HCV full-length or subgenomes with increased replication efficiency. Both constructs with a selectable resistance gene and constructs without one or with a non-selectable reporter gene (e.g. luciferase) can be produced in this way, since replication of cell culture adapted HCV-RNA constructs can also be demonstrated in non-selected cells due to their high replication efficiency.

10

The cell culture adapted mutants of a HCV-RNA construct or

HCV-RNA full-length genome or HCV subgenome with high
replication efficiency compared to the original HCV-RNA
construct or the original HCV full-length genome are
characterized in that they are obtainable by a method in
which the type and number of mutations in a cell culture
adapted HCV-RNA construct are determined through sequence
analysis and sequence comparison and these mutations are
introduced into a HCV-RNA construct, particularly a HCV-RNA
construct according to one of the claims 4 to 19, or into
an (isolated) HCV-RNA full-length genome, either by sitedirected mutagenesis, or by exchange of DNA fragments
containing the relevant mutations.

A group of preferred HCV-RNA constructs, HCV full-length genomes and HCV subgenomes with high and very high replication efficiency, which are consequently highly suitable for practical use is characterised in that it contains one, several or all of the amino acid or nucleic acid exchanges listed in table 3 and/or one or several of the following amino acid exchanges: 1283 arg -> gly , 1383 glu -> ala , 1577 lys -> arg , 1609 lys -> glu , 1936 pro -> ser , 2163 glu -> gly , 2330 lys -> glu , 2442 ile

-> val. (The numbers refer to the amino acid positions of the polyprotein of the HCV isolate conl, see Table 1).

#### 5 Special features of the nucleotid sequences according to the sequence listings:

#### SEQ ID-NO: 1

Name: I389/Core-3'/wt

- 1.1-341: HCV 5' non-translated region
- 10 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 15 4.1813-10842: HCV Polyprotein from Core up to nonstructural protein 5B
  - 5. 1813-2385: HCV Core Protein; structural protein
  - 6. 2386-2961: envelope protein 1 (E1); structural protein
  - 7. 2962-4050: envelope protein 2 (E2); structural protein
    - 8. 4051-4239: Protein p7
    - 9. 4240-4890: nonstructural protein 2 (NS2); HCV NS2-3 Protease
- 25 10.4891-6783: nonstructural protein 3 (NS3); HCV NS3
  Protease/Helicase
  - 11.6784-6945: nonstructural protein 4A (NS4A); NS3
    Protease cofactor
  - 12.6946-7728: nonstructural protein 4B (NS4B)
- 30 13.7729-9069: nonstructural protein 5A (NS5A)
  - 14.9070-10842: nonstructural protein 5B (NS5B); RNA-dependent RNA-polymerase
  - 15.10846-11076: HCV 3' non-translated region

#### SEQ ID-NO: 2

Name: I337/NS2-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 5 2.342-1181: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1190-1800: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1801-8403: HCV Polyprotein from nonstructural protein 2
  up to nonstructural protein 5B
  - 5. 1801-2451: nonstructural protein 2 (NS2); HCV NS2-3 Protease
  - 6. 2452-4344: nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
  - 7. 4345-4506: nonstructural protein 4A (NS4A); NS3
    Protease cofactor
  - 8. 4507-5289: nonstructural protein 4B (NS4B)
  - 9. 5290-6630: nonstructural protein 5A (NS5A)
- 10.6631-8403: nonstructural protein 5B (NS5B); RNAdependent RNA-polymerase
  - 11.8407-8637: HCV 3' non-translated region

#### SEQ ID-NO: 3

15

25 Name: I389/NS3-3'/wt

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
  - 4.1813-7767: HCV Polyprotein from nonstructural protein 3 up to nonstructural protein 5B

- 5. 1813-3708: nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3 Protease Cofactor
- 7. 3871-4653: Nonstructural protein 4B (NS4B)
  - 8. 4654-5994: Nonstructural protein 5A (NS5A)
  - 9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
  - 10. 7771-8001: HCV 3' non-translated Region

10

25

#### SEQ ID-NO: 4

Name: I337/NS3-3'/wt

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 15 2.342-1181: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1190-1800: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 20 4.1801-7758: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B
  - 5. 1801-3696: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
  - 6. 3697-3858: Nonstructural protein 4A (NS4A); NS3
    Protease Cofactor
  - 7. 3859-4641: Nonstructural protein 4B (NS4B)
  - 8. 4642-5982: Nonstructural protein 5A (NS5A)
  - 9. 5983-7755: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
- 30 10. 7759-7989: HCV 3' non-translated Region

#### SEQ ID-NO: 5

Name: I389/NS2-3'/wt

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from
- encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
  - 4.1813-8418: HCV Polyprotein from Nonstructural protein 2 up to Nonstructural protein 5B
    - 5. 1813-2463: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
    - 6. 2464-4356: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
    - 7. 4357-4518: Nonstructural protein 4A (NS4A); NS3
      Protease Cofactor
- 8. 4519-5301: Nonstructural protein 4B (NS4B)
  - 9. 5302-6642: Nonstructural protein 5A (NS5A)
  - 10.6643-8415: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
  - 11. 8419-8649: HCV 3' non-translated Region

20

10

#### SEQ ID-NO: 6

Name: I389/NS3-3'/9-13F

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 30 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell cultureadapted mutant 9-13F
  - 5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 35 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3

Protease Cofactor . . .

- 7. 3871-4653: Nonstructural protein 4B (NS4B)
- 8. 4654-5994: Nonstructural protein 5A (NS5A)
- 9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

7771-8001: HCV 3' non-translated Region

#### SEQ ID-NO: 7

30

Name: I389/Core-3'/9-13F

- 10 Composition (Nucleotide positions):
  - 1.1-341: HCV 5' non-translated region
  - 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1202-1812: internal ribosome entry site from
- encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
  - 4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adapted mutant 9-13F
    - 5. 1813-2385: HCV Core Protein; structural protein
- 20 6. 2386-2961: envelope protein 1 (E1); structural protein
  - 7. 2962-4050: envelope protein 2 (E2); structural protein
  - 8. 4051-4239: Protein p7
- 9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
  - 10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
  - 11.6784-6945: Nonstructural protein 4A (NS4A); NS3
    Protease Cofactor
  - 12.6946-7728: Nonstructural protein 4B (NS4B)
  - 13.7729-9069: Nonstructural protein 5A (NS5A)
  - 14.9070-10842: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
- 35 15.10846-11076: HCV 3' non-translated Region

#### SEQ ID-NO: 8

Name: I389/NS3-3'/5.1

Composition (Nucleotide positions):

- 1.1-341: HCV 5' non-translated region
- 5 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell cultureadapted mutant 5.1
  - 5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3
  Protease Cofactor
  - 7. 3871-4653: Nonstructural protein 4B (NS4B)
  - 8. 4654-5994: Nonstructural protein 5A (NS5A)
  - 9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

7771-8001: HCV 3' non-translated Region

#### SEQ ID-NO: 9

20

Name: I389/Core-3'/5.1

- 25 Composition (Nucleotide positions):
  - 1.1-341: HCV 5' non-translated region
  - 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from
   encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
  - 4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adapted mutant 5.1
    - 5. 1813-2385: HCV Core Protein; structural protein

- 6. 2386-2961: envelope protein 1 (E1); structural protein
- 7. 2962-4050: envelope protein 2 (E2); structural protein
- 5 8. 4051-4239: Protein p7
  - 9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
  - 10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
- 10 11.6784-6945: Nonstructural protein 4A (NS4A); NS3
  Protease Cofactor
  - 12.6946-7728: Nonstructural protein 4B (NS4B)
  - 13.7729-9069: Nonstructural protein 5A (NS5A)
  - 14.9070-10842: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase
  - 15.10846-11076: HCV 3' non-translated Region

#### SEQ ID-NO: 10

15

Name: I389/NS3-3'/19

- 20 Composition (Nucleotide positions):
  - 1.1-341: HCV 5' non-translated region
  - 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
  - 3.1202-1812: internal ribosome entry site from
- encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
  - 4.1813-7767: HCV Polyprotein from Nonstructural protein 3 up to Nonstructural protein 5B of the cell culture-adapted mutant 19
- 5. 1813-3708: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
  - 6. 3709-3870: Nonstructural protein 4A (NS4A); NS3
    Protease Cofactor
  - 7. 3871-4653: Nonstructural protein 4B (NS4B)
- 35 8. 4654-5994: Nonstructural protein 5A (NS5A)

9. 5995-7767: Nonstructural protein 5B (NS5B); RNA-dependent RNA-Polymerase

7771-8001: HCV 3' non-translated Region

5 SEQ ID-NO: 11

10

20

25

Name: 1389/Core-3'/19

- 1.1-341: HCV 5' non-translated region
- 2.342-1193: HCV Core Protein-Neomycin Phosphotransferase fusion protein; selectable Marker
- 3.1202-1812: internal ribosome entry site from encephalomyokarditis virus; directs translation of the downstream located HCV open reading frame
- 4.1813-10842: HCV Polyprotein from Core up to Nonstructural protein 5B of the cell culture-adpated mutant 19
  - 5. 1813-2385: HCV Core Protein; structural protein
  - 6. 2386-2961: envelope protein 1 (E1); structural protein
  - 7. 2962-4050: envelope protein 2 (E2); structural protein
  - 8. 4051-4239: Protein p7
  - 9. 4240-4890: Nonstructural protein 2 (NS2); HCV NS2-3 Protease
  - 10.4891-6783: Nonstructural protein 3 (NS3); HCV NS3 Protease/Helicase
  - 11.6784-6945: Nonstructural protein 4A (NS4A); NS3
    Protease Cofactor
  - 12.6946-7728: Nonstructural protein 4B (NS4B)
  - 13.7729-9069: Nonstructural protein 5A (NS5A)
- 30 14.9070-10842: Nonstructural protein 5B (NS5B); RNAdependent RNA-Polymerase
  - 15.10846-11076: HCV 3' non-translated Region

The invention is described in detail in the following by way of examples of embodiments and respective diagrams. These are as follows

Fig. 1 A: The structure of a HCV-RNA construct according to the invention

10

15

20

25

30

On the top a diagram of the structure of the complete parental HCV genome is given with the positions of the genes for the cleavage products core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B within the polyprotein, and the 5' and 3' non-translated regions (5' NTR and 3' NTR) - shown as a thick horizontal line -, and with the two positions selected for the creation of the subgenomic constructs, namely the position of the 'GDD catalytic domain' of the NS5B RNA polymerase (GDD) and the position of the 3' end of the HCV-IRES (nucleotide positions 1 to 377 and 1 to 389) - drawn above the diagram of the genome. The numbers below the diagram of the genome indicate the respective nucleotide positions.

Diagrams of the structure of two modified HCV-RNA constructs (subgenome) according to the invention are shown below, consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (Neo<sup>R</sup>), the EMCV-IRES (E-I) and the HCV sequences of NS2 or NS3 up to the authentic 3' end. The position of the 10-amino acid deletion comprising the NS5B polymerase GDD motive is marked with a triangle ( $\Delta$ ).

Fig. 1 B: The result of a denaturing formaldehyde-agarose gel electrophoresis for the detection of replicating plus-strand RNA in transfected subpassaged Huh 7 cell clones.

The positions of HCV specific RNAs (arrows) and the 28S rRNA are specified to the right of lane 12, the size (number of nucleotides) of the RNA marker (M) is specified to the left of lane 1.

Fig. 1 C: The result of a PCR test with subsequent Southern Blot to demonstrate the absence of integrated replicon DNA in most of the selected cell clones.

5

10

15

20

25

30

The lanes 1 and 2 show the positive control, lane 13 the negative control. The figures to the left of lane 1 indicate the size of the nucleotide marker molecules.

Fig. 2 A: The result of a PCR test with subsequent Southern Blot for the detection of the sensitive exclusion of integrated replicon DNA (plasmid molecule  $I_{377}/NS3-3^{\dagger}/wt$ ) in a cell clone containing a HCV-RNA construct (9-13).

The lanes 7 to 11 represent the result of a titration of DNA molecules of the construct  $I_{377}/NS3-3$ '/wt without addition of total DNA of the cell clone 9-13, and the lanes 2 - 6 represent the result obtained with the same plasmid molecules with the addition of 1  $\mu$ g 9-13 DNA each prior to the PCR (for the purpose of excluding an inhibitor of the PCR in the DNA preparation). Lane 13 represents the negative control (PCR without DNA template). Lane 1 shows the result, which was achieved with one  $\mu$ g total DNA of the cell clone 9-13.

35 Fig. 2 B: The result of a Northern Blot test to quantify the amounts of HCV plus- and minus-strand RNA.

The arrows mark the positions of replicon RNA.

The "plus" and "minus" details indicate the positive (plus) and negative (minus) polarity of the RNA controls, which have been applied to the gel. "Minus-strand" and "Plus-strand" indicate the specificity of the radioactive RNA probes.

Fig. 2 C: The result of a formaldehyde-agarose gel electrophoresis after radioactive labeling of the intracellular replicated HCV-RNA to demonstrate the resistance of HCV-RNA replication to dactinomycin.

5

30

35

- 15 Fig. 3 A: The detection of HCV specific antigens in the selected cell clones by means of immunoprecipitation after metabolic radioactive labeling.
- The lanes 7 9 represent authentic size marker (which have been obtained in Huh 7 cells after the transient expression of a HCV-RNA construct); identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified to the right of lane 9.
  - Fig. 3 B: Results of an immunofluorescence test to establish the subcellular localisation of HCV antigens.
  - Fig. 4: Diagram of the structure of a selectable HCV-RNA construct according to the invention (complete genome) consisting of the 5' HCV-IRES, the neomycin phosphotransferase gene (NeoR), a heterologous IRES element, for example from the encephalomyocarditis virus (E-I), the complete HCV open reading frame and the authentic 3' NTR.

Fig. 5: Diagram of the structure of HCV-RNA constructs with inserted antibiotic resistance gene (A) within the nucleotide sequence encoding the polyprotein (monocistronic RNA construct), and (B) within the 3' NTR (bicistronic RNA construct).

5

- Fig. 6: Diagram of the structure of HCV-RNA constructs with inserted reporter genes (A) as part of a HCV 10 replicon from NS3 to NS5B, - in the end the reporter protein is cleaved by viral or cellular proteases out of the polyprotein or the marker for selection (selection gene) 15 resistance gene are transfected into the cells through co-transfection -, (B) as part of a fusion gene composed of a resistance and reporter gene (for example for the neomycin phosphotransferase and green fluorescent 20 protein), (C) as part of a replicon composed of a resistance and reporter gene (for example the neomycin phosphotransferase and green fluorescent protein) bound via a nucleotide sequence, which encodes for an amino acid sequence (area with hatches), which can be cleaved by a protease or 25 has a self-cleaving (autocatalytic) activity, (D) as independent gene (in this case the green fluorescent protein), which is expressed from its own internal ribosome binding site (IRES); - the 30 resistance gene (in this case the neomycin phosphotransferase gene) is also expressed from its own internal ribosome binding site (IRES) (polycistronic construct).
- 35 Fig. 7: Diagram of the structure of a HCV-RNA construct, where the resistance gene is linked to the HCV-

RNA sequence via a ribozyme or a recognition site for a ribozyme.

The thick lines illustrate the HCV 5' and 3' NTRs, E-I is a heterologous internal ribosome binding site, which is required for the expression of the resistance gene, and the grey square illustrates the ribozyme or a recognition site for a ribozyme.

10

5

Fig. 8: Diagram of the structure of a HCV-RNA construct with resistance gene and integrated foreign gene.

Fig. 9: Method for comparing the specific infectivity 15 (expressed as number of cell colonies obtained) of total RNA against in vitro transcripts. HCV-RNA is generated by in vitro transcription of a corresponding RNA construct and quantified by measurement of the optical density at 260 nm (OD 20 260 nm). A defined number of these molecules is mixed with a specified amount of total RNA from naive Huh-7 cells and this mixture is transfected naive Huh-7 cells with into the electroporation. At the same time the total RNA 25 of a cell clone, produced by the method described in Figure 1, is isolated using a known state of method art and the amount of HCV-RNA contained therein is determined by means of Northern Blot using a HCV specific RNA probe and 30 subsequent quantification via phosphoimaging. A defined amount of this total RNA is analogously transfected in naïve Huh-7 cells. These cells from both the cultures are then subjected to a G418 selection and the number of colonies created 35 determined by counting after fixing staining with Coomassie Brilliant Blue. For the determination of transfection efficiency lug of a

plasmid allowing the expression of luciferase is added to each transfection culture. An aliquot of the transfected cells is collected after 24 hours and the luciferase activity determined in the respective cell lysates. The number of colonies is always normed to the luciferase expression.

Fig. 10: Sequence analysis of the 9-13 clone. Total RNA of the 9-13 cell clone, resulting from transfection of HCV-RNA construct I377/NS3-3', was isolated 10 using a known state of the art method and the HCV-RNA construct amplified from nucleotide position 59 to 9386 with the aid of 'longdistance RT-PCR' using primer S59 and A9413. The PCR fragments were cloned and 11 clones (9-13 A -15 K) completely sequenced. Clones D and I, E and G as well as H and J turned out to be identical, respectively. The positions of the amino acid differences in the NS3-5B region between the 20 recloned HCV-RNAs and the parental construct are marked with a thick vertical mark for each clone. Each clone was digested with restriction enzyme Sfil and the respective fragment inserted in the parental construct. These clones were each Huh-7 cells transfected in and the cells 25 subjected to selection as described in Figure 1. The number of cell clones obtained with each construct is noted next to the respective construct on the right.

30

35

5

Fig.11 A: Principle of determination of replication with the aid of a reporter gene. In the upper part of the figure, the HCV-DNA construct I<sub>389</sub>/Luc/NS3-3' is shown. It consists of the HCV 5' NTR (nucleotide positions 1-389), the luciferase gene (luc), the IRES of the encephalomyocarditis virus, the HCV NS3-5B and the 3' NTR. The

5

10

30

35

position of the active centre of the NS5B RNA polymerase, into which a deactivating amino acid exchange was introduced, is indicated by 'GND'. plasmids. which code for the construct which is able to replicate or restriction defective, are digested with the an in vitro Scal and added to enzyme transcription reaction with T7 RNA polymerase. After removal of the template DNA, the respective HCV-RNA constructs were transfected in naive Huh-7 cells by means of electroporation and the latter collected at regular intervals.

- in Fig.11 B: Comparison of luciferase activity transfected with the parental HCV-RNA construct 15  $I_{389}/Luc/NS3-3'/wt$  (wt) or the following variants: inactive RNA (318 DN), variants 9-13F or variant 5.1. The cells were collected 6 (not shown), 24, 168 hours after 120, 144 and 72, 96, transfection and luciferase activities determined 20 by luminometric measurement.
- Fig. 12: Selectable HCV full-length genomes (constructs I389/core 3'/5.1 and I389/core 3'/9-13F).

  (A) Diagram of the full-length construct. The region between both indicated recognition sites for the restriction enzyme SfiI corresponds to the sequences of the highly adapted RNA variants 5.1. or 9-13F.
  - (B) Number of colonies which were obtained after transfection of 0.1  $\mu g$  in vitro transcribed RNA of the construct  $I_{389}/core-3$  '/5.1. described under A into HUH7-cells. The result of a representative experiment is given.
  - (C) Demonstration of autonomously replicating HCV full-length RNAs in G418 resistant cell clones which were obtained after transfection of the

5

10

15

transcript. corresponding in vitro autoradiogram the of illustration shows Northern Blot, which was hybridised with a probe against the neo-resistance gene and the HCV 5' NTR. The controls shown in lanes 1 and 2 each correspond to 108 molecules of the indicated in vitro transcripts, mixed with total RNA from naive Huh-7 cells. The negative control contains only total RNA from naive Huh-7 cells (lane 3). Lanes 4-9 contain 3-10 µg total RNA from G418 resistant cell clones, which were obtained after transfection by in vitro transcribed I389/core  $3^{5}.1$  RNA or  $I_{389}/core$   $3^{9}-13F$  RNA. The G418 concentration used for the selection is indicated in each case. Five of the cell clones shown contain the highly adapted RNA variant 5.1 (lanes 4-8), one contains the adapted RNA variant 9-13F (lane 9).

- Fig. 13: HCV-RNA constructs with a reporter gene. 20 Bicistronic HCV-RNA constructs. The reporter gene is translated with the aid of a separate IRES. constructs. HCV-RNA (B) Monocistronic reporter gene product is expressed as fusion protein with a HCV protein. Both portions are 25 linked via a recognition sequence for a viral or cellular protease, which permits a proteolytic separation of the two fused protein portions. In the example shown the reporter gene product and respective HCV protein was fused through a 30 recognition sequence for ubiquitin (Ub).
  - Fig. 14: Tricistronic full-length HCV-RNA construct, that in addition to the resistance gene possesses an inserted foreign gene (ther. gene).

Fig. 15: Monocistronic HCV-RNA constructs, for which the resistance gene product is expressed as a fusion protein with HCV portion. The resistance gene (RG) is either active as a fusion protein or it is fused with the HCV portion via a proteolytically cleavable sequence in such a way that the resistance gene product is split from the HCV portion by a cellular or viral protease. In the example shown the resistance gene was fused with the respective HCV portion through the sequence coding for ubiquitin (Ub).

#### 15 Example 1: Production of HCV-RNA constructs

5

10

- (A) Synthesis and cloning of a complete HCV consensus genome by means of RT-PCR
- The HCV genome, which means the HCV-RNA, was isolated from the liver of a chronically infected patient as described in the following:
- The total RNA was isolated from approx. 100 mg of liver according to the method described by Chomczynski and Sacci (1987, Anal. Biochem. 162, 156). Using 1 µg of this isolated RNA a reverse transcription was carried out with primers A6103 (GCTATCAGCCGGTTCATCCACTGC) or A9413 expand reverse CCTGGAG) and the (CAGGATGGCCTATTGG (Boehringer Mannheim, Germany) system transcriptase 30 manufacturer's recommendations. the according to polymerase chain reaction (PCR) was carried out with the products of this reverse transcription using the expand long template system (Boehringer Mannheim, Germany), in which the buffer containing 2 % of dimethyl sulfoxide was used. After one hour at 42 °C, 1/8 of the reaction mixture was used for a first PCR with primers A6103 and S59

(TGTCTTCACGCAGAAAGCGTCTAG) or A9413 and S4542 (GATGAGCT CGCCGCGAAGCTGTCC). After 40 cycles, 1/10 of this reaction mixture was used for a second PCR with primers S59 and S4542 (AGCACAGCCCGCGTCATAGCACTCG) or (TTAGCTCCCCG TTCATCGGTTGG. After 30 cycles the PCR products means of preparative agarose purified by electrophoresis and the eluted fragments were ligated into the vector pCR2.1 (Invitrogen) or pBSK II (Stratagene). Four clones of each fragment were analyzed and sequenced, and a consensus sequence was established. For this purpose DNA sequences were compared to each other. The positions, where the sequence of one fragment was different from the others, were considered as undesired mutations. In sequence, ambiguities of the case of overlapping PCR fragments of the respective region were amplified and several clones sequenced. By this means several potential mutations could be identified in each fragment and consequently an isolate specific consensus sequence could be established. This established consensus sequence or genome belongs to the worldwide spread genotype 1b. The non-translated region at the 3' end (=3' NTR) was obtained by means of a conventional PCR, whereby an antisense primer, which covers the last 24 nucleotides of the 'X-tail' known from the state of the art was used (Tanaka et al., 1995, Biochem. Biophys. Res. Commun. 215, 25 Rice, PCT/US 96/14033). The authentic und translated region on the 5' end (=5' NTR) downstream of the T7 promoter was created by means of PCR, whereby an which corresponds to a oligonucleotide, promoter (TAA TAC GAC TCA CTA TAG) and the first 30 nucleotides of HCV, was used on one hand and one of the previously mentioned plasmids, which carries one of the 5' fragments of the genome, was used on the other hand. A complete HCV consensus genome was assembled from the subgenomic fragments with the smallest number of non-35 consensus replacements and inserted into a modified pBR322 vector. Deviations from the consensus sequence

eliminated by means of site-directed mutagenesis. In order to produce run-off transcripts with an authentic 3' end, the 3' NTR of the isolates (with the end TGT) was modified to AGT (according to the sequence of the genotype 3 = clone 5 'WS' according to Kolykhalov et al., 1996, J. Virol. 70, 3363) and an additional nucleotide replacement was carried out at position 9562, to preserve the A:T pairing in the hairpin structure at the 3' end of the 3' NTR (Kolyhalov et al. ibid.). In order to eliminate an internal restriction site for the Scal enzyme, a silent nucleotide replacement was further carried out. After joining the full-length genome with the appropriate 5' and 3' NTRs the complete HCV sequence was analyzed. No undesired nucleotide replacement was detected.

15

10

The HCV genome produced in this way should be hepatotropic according to the definition.

#### (B) Synthesis of selectable HCV subgenomic constructs

20

25

30

35

By means of the consensus genome described under (A), HCV subgenomic constructs were created, which include the antibiotic resistance gene neomycin phosphotransferase (NPT) and two sequences of internal ribosome entry sites (IRES). The biochemical procedures used for this are known and familiar to the expert (see: Sambrook, J., Fritsch, T. Maniatis, 1989, Molecular cloning: a laboratory 2nd ed., Cold Spring Harbour Laboratory, Spring Harbor, N.Y.; Ausubel et al. (eds.), 1994, Current Protocols in Molecular Biology, Vol. 1-3, John Wiley & Sons New York). The antibiotic resistance inserted immediately after the 5' NTR, obtaining a bicistronic RNA as a result (see Fig. 1 A). However, the antibiotic resistance gene might as well be inserted at another site of the HCV subgenomic construct, for example encoding nucleotide sequence polyprotein, obtaining a monocistronic RNA as a result (see

Fig. 5 A), or in the 3' NTR (see Fig. 5 B). The IRES elements are one of the two HCV-IRES variants from nucleotides 1-377 or nucleotides 1-389 as well as the IRES of the encephalomyocarditis virus, which controls the translation of the HCV sequence downstream from the genes for NS2 or NS3 up to the authentic 3' end of the genome.

The two HCV-IRES variants mentioned were determined in the following way:

10

Based on deletion analyses of the 3' end of the HCV-IRES (Reynolds et al.1995, EMBO J. 14, 6010) various segments of the 5' NTR were fused with the NPT gene and analyzed with regard to the maximum number of colonies formed by means of co-transfection with a plasmid containing the T7 RNA polymerase gene. The best results were achieved with the HCV sequences from 1-377 to 1-389. Due to the fact that the AUG start codon of the HCV polyprotein is located at position 342 and is consequently included in the IRES sequence, 12 or 16 amino acids of the HCV core protein 20 fused with the protein") were phosphotransferase (see Fig. 1 A).

Accordingly, these modified HCV subgenomic constructs were given the designations  $I_{377}/NS2-3'$  (or  $I_{377}/NS3-3'$ ) and  $I_{389}/NS2-3'$  (or  $I_{389}/NS3-3'$ ). They are illustrated in a diagram in Fig. 1A.

Different cell lines and primary cell cultures of human hepatocytes were transfected with in vitro transcripts of these modified parental HCV subgenomic constructs  $I_{377}/NS2-3'$  (or  $I_{377}/NS3-3'$ ) and  $I_{389}/NS2-3'$  (or  $I_{389}/NS3-3'$ ).

As a parallel negative control for all transfection 35 experiments, an appropriately modified, but defective subgenome was constructed for every modified parental HCV subgenomic construct, which differs from the parental construct due to the fact, that it has a deletion of 10 amino acids within the reading frame comprising the active site of the NS5B RNA polymerase (Behrens et al., 1996, EMBO J. 15, 12; and Lohmann et al., 1997, J. Virol. 71, 8416).

#### (C) Synthesis of selectable HCV genomic constructs

5

20

30

A NS2-3' subgenomic construct, linked at its 5' end to a fragment of the luciferase gene and the complete EMCV-IRES, was restricted with NcoI and SpeI and purified using gel electrophoresis. preparative agarose obtained in this way was ligated with a NcoI/NotI-HCV fragment corresponding to the nucleotide positions 342 to genome and a NotI/SpeI fragment HCV the 1968 corresponding to the nucleotide positions 1968-9605 in a 3resulting construct, where The factor-ligation. complete HCV open reading frame and the 3' lie downstream from the luciferase gene fragment and the EMCV-IRES, was then restricted with PmeI and SpeI and ligated with the analogously restricted I389/NS3-3'/wt subgenomic construct vector. This selectable HCV genomic construct is illustrated in Fig. 4.

# (D) Production of in vitro transcripts corresponding to the HCV-RNA constructs

previously described purified plasmid DNAs linearized with ScaI and used for an in vitro transcription reaction after phenol/chloroform extraction and isopropanol precipitation using the following components: 80 mM HEPES,  $MgCl_2$ , Spermidine, 2 mM Нq 7.5, 12,5 mM Dithiothreitol, 2 mM of each NTP, 1 unit RNasin/µl, 50 µg/ml restricted DNA and approx. 2 units/µl polymerase. After 2 hrs. at 37 °C half of the amount of T7 polymerase was added and the reaction mixture was incubated for two further hours. In order to remove DNA the mixture was extracted with acid phenole (U. Kedzierski, J.C. Porte,

precipitated Techniques 10, 210), Bio 1991, the pellet was dissolved in water isopropanol, incubated with DNase (2 units per µg DNA) for 60 min. at 37 °C. After subsequent extraction with acid phenole, acid phenole/chloroform and chloroform as well as isopropanol precipitation the dissolved RNA was quantified using optical density measurement and its integrity was checked using formaldehyde-agarose gel electrophoresis.

10

### Example 2: Transfection experiments with the hepatoma cell line Huh 7

With all transfection experiments it was carefully ensured that any template DNA had been removed beforehand so as to avoid the possibility of this DNA integrating in transfected cells and conferring a neomycin resistance to them independent from HCV replication. The reaction mixture was therefore treated with 2 units of DNase per µg DNA for 60 min. at 37 °C after the in vitro transcription and extracted with acid phenole, acid phenole/chloroform and chloroform. Prior to being used for the transfection the precipitated RNA was analyzed using formaldehyde-agarose gel electrophoresis.

25

30

Three separate transfection experiments were carried out with the highly differentiated human hepatoma cell line Huh 7 (according to Nakabayashi et al. 1982, Cancer Res. 42, 3858). Each time 15 µg RNA were introduced into 8 x 10<sup>6</sup> Huh 7 cells by electroporation and the cells were seeded in culture dishes with a diameter of 10 cm. 24 hours after seeding, neomycin (= G418) was added in a final concentration of 1 mg/ml. The culture medium was changed twice a week. After 3 - 5 weeks small colonies were visible, which were isolated and grown under the same culture conditions.

The cell clones obtained during the first experiment were isolated and subpassaged. Most of the clones died during this procedure, and the final yield was only 9 clones from cells, which had been transfected with the parental HCV subgenomic constructs and 1 clone (clone 8-1) from cells, which had been transfected with a defective HCV genomic construct, namely a defective NS2-3' HCV-RNA. Apart from an extended doubling time and the occasional occurrence of irregularly shaped cells, no consistent morphological differences were found between these 9 cell clones and the single cell clone (clone 8-1) or the parental Huh 7 cells.

The main criteria for functioning HCV genomic constructs are the formation of viral RNA with the correct size and the absence of (integrated) plasmid DNA, which could transfer or mediate on a G418 resistance.

To determine the HCV-RNA in the Huh 7 cells, the total RNA was isolated and analyzed by means of the common Northern Blot method using a plus-strand specific ribo probe (RNA probe). For this purpose the total RNA was isolated from respective cell clones according to the described by Chomczynski and Sacchi 1987, Anal. Biochem. 162, 156 and 10 µg RNA, which is equivalent to the total RNA content of  $0.5 - 1 \times 10^6$  cells, are separated using denatured formaldehyde-agarose gel electrophoresis (lanes 3 to 12 of Fig. 1 B). At the same time 109 in vitro transcripts (ivtr.), which correspond to the I389/NS2-3'/wt or the  $I_{389}/NS3-3'/wt$  replicon RNAs, are separated as well as size markers with authentic sequence (lane 1 or lane 2). 30 transferred nylon-based separated RNA was onto membranes and hybridized with a radioactively labeled plusstrand specific RNA probe, which was complementary to the complete NPT gene and the HCV-IRES of nucleotide 377 to nucleotide 1. The positions of the HCV specific RNAs 35 (arrows) and the 28S rRNA are specified to the right of lane 12, the size (amount of nucleotides) of the RNA marker

is specified to left of lane 1. The RNA marker fragments contain HCV sequences and therefore hybridize with the ribo probe (= RNA probe). The results of this analysis are illustrated in Fig. 1 B.

5

15

20

25

30

35

With the exception of clone 8-1 transfected with the defective HCV genomic construct, all cell clones produced homogenous HCV-RNAs of correct lengths (approx. nucleotides in the case of the NS2-3' and approx. 7970 nucleotides in the case of NS3-3' replicon). This result is an indication of the fact, that the functional replicons or the functional HCV genomic constructs transfer the G418 resistance. In order to exclude, that the G418 resistance is caused by a plasmid DNA, which is integrated in the genome of the Huh 7 host cell and transcribed under control of a cellular promoter, the DNA of each clone was analyzed by means of a NPT gene specific PCR. Consequently the DNA was isolated from the selected Huh 7 cell clones by means of digestion with proteinase K (40 $\mu$ g/ml, 1h, 37 °C) in 1mM EDTA, 0,5 % SDS and subsequent 10mMTris, pH7,5, extraction with phenol, phenol/chloroform and isopropanol precipitation. The DNA precipitate was dissolved in 10 mM Tris (pH 7,5) and 1 mM EDTA and incubated with Rnase A for Following a phenol/chloroform extraction and ethanol precipitation, 1  $\mu g$  DNA, equivalent to 4 - 8  $\times$  10<sup>4</sup> cells, was analyzed by means of PCR using NPT gene specific (5'-TCAAGACCGACCTG TCCGGTGCCC-3' and primers CTTGAGCCTGGCGAACAGTTCGGC-3'), and a DNA fragment consisting of 379 nucleotides was generated. The specificity of the PCR product was established by means of the Southern Blot method, in which a DNA fragment labeled by digoxigenin was used, which corresponds to the NPT gene. As positive controls (for the detection of possibly contaminating nucleic acids) the PCR method was carried out with 107 plasmid molecules or 1 µg DNA from a BHK cell line, which was stably transfected with a neomycin resistance gene, and . . .

as negative control, the PCR was carried out with the same reagents but without added DNA.

The results of this analysis are illustrated in Fig. 1 C. The lanes 1 and 2 represent the positive controls, lane 13 represents the negative control. The numbers to the left of lane 1 indicate the sizes of the nucleotide marker molecules.

A NPT-DNA could not be detected in any cell clone, apart 10 from clone 7-3 (Fig.1C, lane 3), which was obtained from cells after transfection with a NS2-3' replicon/NS2-3'HCV genomic construct, and clone 8-1 (Fig.1C, lane 12), which was obtained from cells after transfection with a defective HCV genomic construct. This result was another indication 15 of the fact, that the G418 resistance of most clones was passed on by the replicating HCV-RNA. But even regardless of these results, it is unlikely, that integrated plasmid correct sizes, because produces HCV-RNAs of plasmids used for the in vitro transcription contain eukaryotic promoter nor a polyadenylation neither an the case of clone 7-3 the resistance is signal. In therefore very probably passed on by the HCV-RNA construct or the replicating HCV-RNA as well as by an integrated NPT DNA sequence, whereas the resistance of the cells of clone 25 8-1 is only caused by the integrated plasmid DNA.

Clone 9-13 (Fig. 1 B, lane 11) was subject to further tests, to confirm that the G418 resistance was passed on by a self-replicating HCV-RNA. Clone 8-1, which carries the integrated copies of the NPT gene, was used throughout as negative control. A PCR was carried out, which allows the detection of < 1000 NPT gene copies in ~ 40.000 cells, with the aim to rigorously exclude the presence of NPT-DNA in clone 9-13. The result of this PCR is illustrated in Fig. 2A. The PCR proceeded in detail as follows:

30

35

During the test,  $10^6$  -  $10^2$  plasmid molecules ( $I_{377}/NS3-3'/wt$ ) were used either directly (lanes 7 - 11) or after adding each 1 µg 9-13 DNA (lanes 2 - 6). The specificity of the amplified DNA fragment was determined by Southern Blot using a NPT specific probe. A PCR without DNA probe was carried out as negative control (lane 12).

Even with this sensitive method, no plasmid DNA could be detected in one  $\mu g$  DNA of the cell clone 9-13 (lane 1). To estimate the amount of HCV plus- and minus-strand RNAs in these cells, a dilution series of total RNA was analyzed with the Northern Blot method using a plus- or minus-strand specific radioactively labeled ribo probe (= RNA probe). For this purpose, 8, 4 or 2 µg of total RNA, which have been isolated from the cell clones 9-13 and 8-1, analyzed in parallel to known amounts of in vitro transcripts with plus- or minus-strand polarity (control RNAs) in the Northern Blot method and were then subjected to a hybridization. The hybridization was carried out with which covered the a plus-strand specific ribo probe, complete NPT gene and the HCV-IRES ('plus-strand', panel), or with a minus-strand specific RNA probe, which was complementary to the NS3 sequence ('minus-strand', bottom panel). The arrows mark the positions of replicon RNA. The results of this analysis are illustrated in Fig. 2 B.

Approx. 10<sup>8</sup> copies/µg total RNA were detected in the case of the plus-strand, which is equivalent to 1000 - 5000 HCV-RNA molecules per cell, whereas the amount of minus-strand RNA was 5- to 10-times less. This result corresponds to the assumption, that the minus-strand RNA is the replicative intermediate form or intermediate copy, which is used as a template for the synthesis of the plus-strand molecules.

35

10

15

20

25

Due to the fact that the reaction is mainly catalyzed by the viral RNA-dependent RNA polymerase, the synthesis of

the HCV-RNAs should be resistant to dactinomycin, antibiotic, which selectively inhibits the RNA synthesis of templates, but not the RNA synthesis from templates. To confirm this assumption, cells were incubated with [3H] uridine in the presence of dactinomycin, the radioactively labeled RNAs were extracted, separated by electrophoresis and means of denaturing agarose gel analyzed with the aid of a common Bio-Imager using a [3H]sensitive screen. For this purpose approx.  $5 \times 10^5$  cells of the clones 9-13 and 8-1 were incubated at a time with 100  $\mu$ 10 Ci [3H] uridine for 16 hrs. in the absence (-) or presence (+) of 4  $\mu$ g/ml of dactinomycin (Dact). Following this labeling reaction the total RNA was prepared and analyzed by means of formaldehyde-agarose gel electrophoresis. Only 1/10 of the total RNA is illustrated in the first two 15 lanes. The radioactively labeled RNA was visualized using a BAS-2500 Bio-Imager (Fuji).

The results of this analysis are illustrated in Fig. 2 C. In accordance with the inhibitor profile of the NS5B polymerase (Behrens et al., 1996, EMBOJ. 15, 12 and Lohmann et al., 1997, J. Virol. 71, 8416), dactinomycin had no influence on the replication of the HCV RNA, whereas the synthesis of cellular RNA was inhibited. A RT-PCR was carried out for recloning the replicating sequences, to confirm the identity of the viral RNA. The sequence analysis of the recloned RNA showed that the RNA in clone 9-13 is HCV specific and corresponds to the transfected transcript of the HCV construct I<sub>377</sub>/NS3-3'/wt.

30

20

For the analysis of the viral proteins, at first the respective cells were metabolically radioactively labeled with [35S] methionine/cysteine, then lysed and afterwards the HCV specific proteins were isolated from the cell lysates by means of immunoprecipitation. The results of these analyses are illustrated in Fig. 3 A. The detailed procedure was as follows: Cells of the cell clones 9-13

(wt) and 8-1 ( $\Delta$ ) were metabolically radioactively labeled by treating them with a protein labeling mixture familiar to the expert and available on the market (for example. NEN Life Science). The HCV specific proteins were separated from the cell lysate by immunoprecipitation (IP) under nonaccording denaturing conditions (for example Bartenschlager et al., 1995, J. Virol. 69, 7519) using three different antisera (3/4, 5A, 5B, according to the labeling on the top end of the lanes 1 to 12). The immune complexes were analyzed by means of tricine SDS-PAGE and made visible by means of autoradiography. To authentic size markers, the homologous replicon construct  $I_{377}/NS3-3'/wt$  was subject to a transient expression by the vaccinia virus-T7 hybrid system in the Huh 7 cells. The resulting products were used as size markers (lanes 7 - 9) parallel to the cells of the clones 9-13 and 8-1. Identified HCV proteins are labeled on the left edge of lane 1, the molecular weights (in Kilodalton) are specified. on the right edge of lane 9. It should be noted that the NS3/4 specific antiserum ('3/4') used preferably reacts with NS4A and NS4B causing an underrepresentation of NS3.

10

All viral antigens could unambiguously be detected, and apparent molecular weights did not show their detected after transient those being difference to expression of the same bicistronic HCV-RNA construct in the original Huh 7 cells. An immunofluorescence detection reaction was carried out using NS3 and NS5A specific antisera, to determine the subcellular distribution of the viral antigens (for example according to Bartenschlager et al., 1995, J. Virol. 69, 7519). For this purpose cells of were fixed with the clones 9-13 (wt) and 8-1 ( $\Delta$ ) methanol/acetone 24 hrs. after incubating on coverslips and incubated with polyclonal NS3 or NS5A specific antisera. The bound antibodies were made visible with a commercially available FITC conjugated anti-rabbit antiserum. The cells

were counterstained with the 'Evans Blue' stain to suppress unspecific fluorescence signals.

The results of this detection test are illustrated in Fig. 3 B. A strong fluorescence in the cytoplasm could be detected with both antisera. The NS5A specific antiserum also caused a slight nuclear fluorescence, which indicates that at least small amounts of this antigen also reach the nucleus. But the generally dominating presence of the viral antigens in the cytoplasm are a strong indication that HCV-RNA replication occurs in the cytoplasm, as is the case with most RNA viruses.

These results prove clearly that the establishment of a cell culture system for the HCV could be accomplished with the test arrangement described, the efficiency of which surpasses everything known up until now by far and for the first time allows the detection of viral nucleic acids and biochemical with conventional and approved proteins efficiency actually. This allows detailed methods. 20 pathogenesis, genetic analyses examination of HCV different HCV functions and a precise study of the virus/host cell interaction, through which new starting points for the development of a antiviral therapy can be 25 defined.

### Example 3: Transfection of Huh 7 cells with HCV total constructs

30

35

Huh7 cells are transfected and selected as described in example 2, whereby in this case selectable constructs are used, which contain the complete virus genome. Corresponding to Example 2, the resulting cell clones are tested for the absence of HCV-DNA by means of PCR and the productive replication of HCV-RNA is then established by means of Northern Blot, [3H]uridine labeling in the

presence of dactinomycin, detection of the viral proteins or antigens preferably with the aid of the Western Blot, the immunoprecipitation or immunofluorescence. In contrast to the arrangements described in Example 2, the construct described here makes it possible to obtain more complete and very likely infectious viruses, which has not been the case in the subgenomic constructs described in Example 2. These viruses existing in the cell and the cell culture supernatant are concentrated for example by means ultracentrifugation, immunoprecipitation or precipitation with polyethyleneglycol, and all exogenous nucleic acids, which means those that are not incorporated into the virus particle, are digested by incubation with nucleases (RNase, DNase, micrococal nuclease). In this way all contaminating nucleic acids, which are not included in the protecting virus particle, can be removed. After inactivation of the nucleases, the protected viral RNA is isolated for example by means of incubation with proteinase K in a buffer by extracting with phenol SDS, containing phenol/chloroform and detected by means of Northern Blot or RT-PCR using HCV specific primers. Also in this arrangement, the combination of the HCV consensus genome described with a selection marker was crucial for the efficient production of viral RNA, viral protein and 25 therefore HCV particles.

15

30

35

application and of Example 4: Production RNA construct, whereby the resistance gene is to the HCV subgenomic sequence via a ribozyme or a recognition site for a ribozyme.

A HCV-RNA construct is produced according to Example 1 or Example 3, where an antibiotic resistance gene is linked to the HCV-RNA sequence through a ribozyme or a recognition sequence for a ribozyme. These constructs are illustrated in a diagram in Fig. 7. Huh 7 cells are transfected with

this HCV-RNA construct as described in Example selection with the appropriate antibiotic follows the transfection into the cells. The inserted ribozyme is activated in the cell clones obtained in the procedure or, in the case of a construct, which carries a recognition sequence for a ribozyme, the ribozyme is transfected into the cell (for example by means of transfection of a ribozyme construct or infection with a viral expression vector, into which the respective ribozyme has been inserted). In both cases the resistance gene is separated from the HCV-RNA sequence by the ribozyme-dependent cleavage. The result in the case of the HCV genome is an authentic HCV genome without a resistance gene, which can form authentic infectious virus particles. A HCV replicon without resistance gene is created in the case of the HCV subgenomic constructs.

10

20

30

35

### Example 5: Co-transfection of a HCV-RNA construct with a separate luciferase transfection construct

A HCV-RNA construct is produced according to Example 1 (A) or Example 3 or Example 4. At the same time a transfection construct is produced, which comprises the luciferase gene, luciferase gene is linked whereby this nucleotide sequence, which encodes a HCV protease (for example NS3 protease) cleavage site, to a second nucleotide sequence, which encodes for another protein or a part of HCV-RNA construct andi transfection protein. construct are transfected into any host cells, preferably hepatoma cells, most preferably Huh 7 cells. This can be realised as described in Example 2. The product of the modified luciferase gene is a luciferase fusion protein, where the luciferase is inactivated due to the fusion with the foreign part. The fusion protein, which contains a recognition sequence for a HCV protease, is cleaved in transfected cells with high HCV replication,

consequently the active form of the luciferase, which can luminometric measurement, identified through released. If the replication of the HCV-RNA construct is inhibited the fusion protein will not be cleaved and no luciferase will be released. The quantitative determination of the luciferase is therefore a measure for the replication of the HCV construct. Instead of the luciferase gene, another reporter gene can just as easily be used, which is modified in the same way, so that its expression depends on the viral replication, although this reporter gene is not part of the HCV construct. A cellular protein, which is deactivated or activated by the HCV proteins or nucleic acid, can also be used as a so called surrogate marker. The expression or activity of this surrogate marker is in this case a measure for the replication of the viral DNA.

10

# Example 6: Production of HCV subgenomic constructs with integrated foreign genes to be used as liver cell specific vector in gene therapy

These recombinant and selectable HCV subgenomic constructs are transfected in transcomplementing helper cell lines, which means in cell lines, which express the missing functions (for example the structural proteins) inducible or constitutive way. Cell clones containing a functional HCV subgenomic construct can be established appropriate selection. The viral structural proteins expressed from the host cell allow the formation of virus particles, into which the RNA of the HCV subgenomic constructs will be transfected. The results are therefore virus-like particles, which contain subgenomic construct according to the invention including the inserted foreign gene and which can transmit this to other cells by means of infection. An example for this construct is illustrated in Fig. 8.

It is also possible to use the HCV subgenomic construct with integrated foreign gene directly as an expression vector. This involves the same method as mentioned previously except that cell lines, which do not express transcomplementing factors, are transfected. In this case the HCV construct is only used as an expression vector.

### 10 Example 7: Production of cell culture adapted HCV-RNA constructs

#### (A) Method of isolation

The following method was used to determine adaptive 15 mutations and to produce cell culture adapted HCV-RNA constructs: cells were transfected with a HCV-RNA construct as described in Examples 1 and 2 and G418-resistant cell clone produced. For the determination of ability to replicate (understood in this context to be the number of 20 resistant cell clones obtained per microgram of transfected HCV-RNA or HCV-RNA construct), the total RNA from one of the cell clones, [the 9-13 clone (Fig. 1 B, lane 11)], was isolated and the quantity of HCV-RNA contained within it was determined by Northern Blot as described in Fig. 2 B. Ten micrograms of the total RNA, containing approx. 109 molecules of HCV-RNA, was then transfected into naive Huh-7 cells using electroporation (Fig. 9). In parallel, 109 in vitro transcripts of the analoguous neo-HCV-RNA, which had been adjusted with 30 isolated total RNA from naive Huh-7 cells to a total RNA quantity of 10 µg, were transfected in naive Huh-7 cells. After selection with G418, the number of cell colonies, expressed as 'colony forming units (cfu) per microgram RNA', was determined in both these cultures. 35 concentration of 500 µg/ml G418 in the selection medium, the number of colonies obtained with the HCV-RNA contained

in **isolated** total RNA from clone 9-13, was approx. 100,000 cfu per microgram HCV-RNA. In contrast, only 30-50 colonies were obtained with the same quantity of in vitro transcribed HCV-RNA. This result confirms that the specific infectivity of the HCV-RNA isolated from the cell clones is approx. 1,000-10,000 times higher than the infectivity of the analoguous in vitro transcripts. The experimental approach is shown in Fig. 9.

With the aid of 'long-distance RT-PCR', the HCV-RNA was 10 amplified from the total RNA of the 9-13 cells, the PCR amplificate was cloned and numerous clones were sequenced. A comparison of the sequences of these recloned RNAs with the sequence of the RNA originally transfected into the naive Huh-7 cells, showed that the recloned RNAs possessed 15 numerous amino acid exchanges distributed over the whole HCV sequence (Fig. 10). Sfil fragments of these recloned mutants were used to replace the analoguous SfiI fragment of the original replicon construct, and RNAs of the respective mutants were transfected in naive Huh-7 cells. After selection with G418 the number of colonies created was determined for each HCV-RNA mutant. While only 30 - 50 colonies per microgram RNA were obtained with the parental RNA the number of colonies was noticeably higher for two of the recloned variants (Fig. 10). In the case of the HCV-RNA 25 constructs 9-13I and 9-13C the specific infectivity was increased to 100 - 1,000 cfu per microgram RNA and for 9-13F replicon it was 1,000 - 10,000 cfu per microgram RNA. These results show that the amino acid exchanges in the analysed NS3-5B regions of the mutants 9-13I, 9-13C and 30 particularly of 9-13F, led to a considerable increase in ability to replicate. In contrast all the other HCV-RNA constructs (9-13 A, B, G, H und K) were no longer able to replicate, they thus contained lethal mutations.

35

In order to answer the question which of the amino acid exchanges in the 9-13F construct led to an increase in

replication, the exchanges were introduced separately or in combination into the parental HCV-RNA construct, and the corresponding RNAs transfected in naive Huh-7 cells. The result of the transfection with these RNAs is summarised in Table 1. From this it is evident that in the present example the high ability to replicate is determined by several mutations. The amino acid exchanges in the HCV-RNA regions NS5A and NS4B make the greatest contribution. The single exchanges in the NS3-Region also make a contribution and perhaps they are synergistic.

10

20

25

30

35

These results confirm that it was through the G418 selection of the cells transfected with the neo-HCV-RNA construct that there was enrichment of those HCV-RNAs having noticeably higher ability to replicate. HCV-RNA constructs with greatly differing replication efficiencies can be selected using the experimental approach described here. The higher the concentration of the antibiotic in the in/on which the HCV-RNA construct selection medium, containing cells are cultivated for selection, the higher be the extent of adaptive mutations and hence replication efficiency of the relevant HCV-RNA constructs, to allow the cells to grow under these conditions. If the carried out using lower antibiotic are selections concentrations, cells can survive and multiply, but the HCV-RNA construct shows a comparatively lower replication efficiency and fewer adaptive mutations.

As has been shown, the 9-13F HCV-RNA construct described so far, which contains several adaptive mutations, had a higher replication efficiency than the parental HCV-RNA. In order to obtain HCV-RNAs with even higher replication in cell culture, the HCV-RNA contained in the total RNA of a selected cell clone was passaged several times in naive Huh-7 cells. The selected 5-15 cell clone, was obtained by transfection with the HCV-RNA construct  $I_{389}/NS3-3'$  (Fig. 1). It largely corresponds to the cell clone 9-13, produced

by transfection with a HCV-RNA construct, having a HCV-IRES shorter by 22 nucleotides (I377/NS3-3'; Fig. micrograms of total RNA, isolated from cell clone 5-15, into Huh-7 cells naive transfected electroporation and the cells subjected to a selection with 1 mg/ml G418. The total RNA from one of the cell clones thus produced was again isolated, transfected into naive Huh-7 cells and selected in the same way. This process was repeated a total of four times. After the fourth passage the total RNA was isolated from a cell clone and the neo-HCV-RNA amplified with the aid of the 'long-distance RT-PCR'. The amplified DNA fragment was digested with the restriction enzyme SfiI and inserted into the restricted parental construct I389/NS3-3'. Over 100 DNA clones were obtained altogether and then analysed by means of restriction digestion. In vitro transcribed RNA of about 80 of these clones was each transfected into naive Huh-7 cells and subjected to a selection with 500mg/ml G418. Of the 80 neo-HCV-RNA variants examined, the great majority proved to be replication defective. However, the specific expressed as 'colony forming units' infectivity, microgram RNA, was noticeably increased in the case of two mutants, 5.1 and 19 (Table 2). Through several passages of the RNA in cell culture it is clear that HCV-RNAs are produced whose replication efficiency due to mutations ("adaptive mutations") is several orders of magnitude higher than the original RNA cloned from patients.

#### (B) Modified method

30

35

Adaptive mutations produced and identified according to (A) can be transferred into a HCV-RNA construct with low ability to replicate. This leads to a huge increase in the replication of this construct. The increase is so great it can be demonstrated that HCV-RNAs transfected into cell culture can replicate even in the absence of selection pressure Fig. 12 shows a comparison of the replication

efficiency of HCV-RNAs, which corresponded either to the starting sequence or to the adaptive sequences 9-13F or 5.1. For the purposes of simple measurement, the neo-gene was removed and replaced by the gene for luciferase. The negative control used was again a HCV-RNA construct that was replication defective due to a deactivating mutation in polymerase. Already 24 hours NS5B RNA noticeable difference is evident transfection a luciferase activity between the defective RNA and the 9-13F or 5.1 constructs, while hardly any difference could be 10 seen between the defective RNA (318 DN) and the parental RNA construct (wt) that possessed no adaptive mutations. During the whole period of observation, the highest luciferase activity, and thus highest replication, was obtained with the 5.1-RNA. These results not only confirm 15 the high replication efficiency of this RNA, but also show that it is possible to create a cell culture system with adapted HCV-RNA constructs for which the presence of a selectable gene is no longer necessary. A summary of the nucleotide and amino acid differences between the starting construct and the mutants 9-13F, 5.1 and 19 is presented in Table 3.

## 25 Example 8: Production of cell culture adapted HCV-RNA full-length genome

30

In the examples 1 to 7 a subgenomic HCV-RNA was used which lacked the whole structural protein region from core up to p7 or even NS2. It will be shown in this example that it is possible to make a HCV full-length genome replicate in cell culture with the aid of an adapted NS3-5B sequence. For this purpose the SfiI fragment of the highly adapted HCVproduced according to Example 7 transferred into a selectable HCV full-length genome (Fig. 12). This HCV genome was transfected into naive Huh-7 cells various subjected selection with and to

concentrations. Depending on the strength of selection (the G418 concentration), a varying large number of cell clones was obtained (Fig. 12 B). By contrast no colonies were obtained with the parental HCV full-length containing no adaptive mutations, as was the case for the negative control, which was replication defective due to a deactivating mutation in the NS5B RNA polymerase. that the thus resulting cell clones autonomously replicating HCV full-length contained an construct, total RNA from several cell clones was isolated and analysed by means of the Northern Blot method. full-length HCV-RNA was clearly detectable in all cell clones (Fig. 12). It is thus clearly confirmed, that with the aid of cell culture adapted HCV sequences it is possible to produce a HCV full-length genome, 15 replicates highly efficiently and autonomously in a cell line, i.e. adapted HCV full-length genomes can also be produced with the system of the invention. Furthermore, as this clone possesses the complete HCV sequence, i.e. also possesses the structural proteins necessary for virus particle formation, it is possible to produce large quantities of infectious virus particles in cell cultures with this system. As a confirmation of these viruses, cellfree cell supernatants carrying a replicating HCV fulllength genome, are added to naive Huh-7 cells and the thus 25 infected cells subjected to selection with G418. Each cell clone growing under these conditions originates from an infected cell. The viruses in the cell culture supernatant of cells possessing a replicating HCV full-length genome can be enriched and purified using various known state of 30 methods such as ultracentrifugation microdialysis. They can then be used for the infection of naive cells. Using this method it is clearly demonstrated that cell culture adapted full-length genomes can be produced with the HCV cell culture system of the invention. These genomes replicate with high efficiency in cells and produce infectious viruses. The latter can be detected by

infection of an experimental animal, preferably a chimpanzee.

### 5 Example 9: Production of HCV full-length constructs and HCV subgenomic constructs with reporter genes

A HCV-RNA construct is produced in which a reporter gene is inserted in place of the antibiotic resistance gene (Fig. 13). Replication can thereby be determined through the quantity or activity of the reporter gene or reporter gene product. The reporter gene is preferably a gene from the the luciferase genes, the group of (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) or the (secreted alkaline phosphatase gene). its products, namely the relevant reporter gene and reporter proteins, can be detected for example using fluorescence, chemiluminescence, colorimetrically or by means of immunological methods (for example, enzyme-linked immunosorbent assay, ELISA).

The reporter gene can be expressed either from a separate IRES or in the form of a fusion protein, which is active either as such or fused with a HCV protein via a poteolytically cleavable amino acid sequence in such a way that the reporter is separated from the HCV protein by cleavage of a cellular or viral (HCV) protease.

30

10

20

Example 10: Production of HCV full-length constructs with integrated foreign genes used as liver cell specific vectors for gene therapy or as expression vectors.

35

The construct (Fig. 14) is transfected in cells and leads to the formation of HCV virus particles that can be used

for the infection of further cells. Since the virus particles have encapsidated RNA with a foreign gene, it can be used in the infected cells for the production of the protein coded by this foreign gene. Cells transfected with the construct also express the foreign gene.

Example 11: Production of monocistronic HCV-RNA constructs in which the resistance gene product is expressed as a fusion protein with the HCV portion.

10

20

25

30

It is an advantage for some tests if the HCV-RNA construct does not possess a heterologous IRES element. Tests of this type are, for example, the determination of interferon resistance. If a cell possessing a HCV-RNA construct is incubated with interferon alpha or beta, a reduction in replication of the HCV-RNA results. In order to explain the mechanism of this effect it is necessary for the HCV-RNA construct not to possess any heterologous otherwise it is not possible to determine whether the interferon mediated inhibition is via inhibition of the HCV replication or inhibition of the heterologous IRES. For this reason constructs are produced for resistance gene is fused with a HCV protein (Fig. 15). Either the fusion protein is active as such or the resistance gene product is linked to a HCV protein via a proteolytically cleavable amino acid sequence in such a way that it is separated from the HCV protein by a cellular or viral (HCV) protease.

Table 1: Specific infectivities (cfu/ $\mu$ g RNA) of HCV RNA constructs with adaptive mutations found with the 9-13F mutant introduced into the parental construct I<sub>389</sub>/NS3-3'/wt

amino acid exchange <sup>1</sup>	HCV protein	cfu/µg RNA <sup>2</sup>
none		30 - 60
1283 arg -> gly	NS3	200 - 250
1383 glu -> ala	NS3	30 - 60
1577 lys -> arg	NS3	30 - 60
1609 lys -> glu	NS3	160 - 300
(1283 arg -> gly + 1383 glu -	NS3	360 - 420
> ala + 1577 lys -> arg +		·
1609 lys -> glu)		·
1936 pro -> ser	NS4B	1000-5000
2163 glu -> gly	NS5A	1000-5000
2330 lys -> glu	NS5A	30 - 60
2442 ile -> val	NS5A	30 - 60
all together		5000

<sup>&</sup>lt;sup>1</sup> amino acid change in the polyprotein of the HCV isolate con 1 (EMBL-gene bank No. AJ238799); amino acids are given in single letter code.

5

 $<sup>^2</sup>$  Colony forming units (number of cell clones) obtained 0 with a selection of 500  $\mu g/ml$  G418.

Table 2: Specific infectivities (cfu/ $\mu$ g RNA) of the parental HCV RNA construct I<sub>389</sub>/NS3-3'/wt and the variants 9-13I, 9-13F, 5.1 and 19.

Transfected RNA variant	cfu/µg RNA¹		
wild type	30 - 50		
9-13 I	100 - 1.000		
9-13 F	1.000 - 10.000		
5.1	50.000 - 100.000		
19	50.000 - 100.000		

 $<sup>^{1}</sup>$  Colony forming units (number of cell clones) obtained with a selection of 500 $\mu$ g/ml G418.

**Table 3:** Nucleotide and amino acid sequence differences between the parental HCV RNA construct  $I_{389}/NS3-3'/wt$  and the mutants 9-13I, 9-13F, 5.1 und 19

Clone	nt-position	nt-exchange	aa-exchange
9-13 I	3685	C > T	P > L
	4933	C > T	T > M
	5249	T > C	
	8486	C > T	-
	8821	G > A	W > stop
	8991	C > G	R > G
	9203	A > G	-
	9313	T > C	F > S
	9346	T > C	V > A
9-13 F	3866	C > T	-
	4188	A > G	R > G
-	4489	A > C	E > A
	4562	G > A	-
	4983	T > C	<b>-</b>
	5071	A > G	K > R
	5166	A > G	K > E
<del></del>	6147	C > T	P > S
	6829	A > G	E > G
	7329	A > G	K > E
	7664	A > G	I > V
	8486	C > T	-
	8991	C > G	R > G
NK5.1	4180	C > T	T > I
	4679	C > T	-
	4682	T > C	-
	5610	C > A	L > I
	6437	A > G	-
	6666	A > G	N > D

·	6842	C > T	-
	6926	C > T	-
	6930	T > C	S > P
	7320	C > T	P > S
	7389	A > G	K > E
NK19	3946	A > G	E > G .
	4078	C > G	A > G
	4180	C > T	T > I
	4682	T > C	-
	5610	C > A	L > I
	5958	A > T	M > L
	6170	T > A	-
	6596	G > A	-
	6598	C > G	A > G
	6833	C > T	-
	6842	C > T	-
	6930	T > C	S > P
	7141	A > G	E > G
	7320	C > T	P > S
	7389	A > G	K > E
· · · · · · · · · · · · · · · · · · ·	7735	G > A	S > N
	<del></del>		

Given are the differences between the nucleotide and amino acid sequences of the parental HCV RNA sequence con 1 (EMBL-gene bank No. AJ238799) and those of the cell culture adapted HCV RNAs. Numbers refer to the nucleotide and amino acid positions of the con 1 isolate. nt, nucleotide; aa, amino acid.

#### SEQUENCE LISTING

```
<110> Bartenschlager, Ralf
<120> Hepatitis C Virus cell culture system
<130> ba-1
<140> 199 15 178.4
<141> 1999-04-03
<160> 11
<170> PatentIn Ver. 2.1
<210> 1
<211> 11076
<212> DNA
<213> Hepatitis C Virus
<400> 1
qccaqcccc qattqqqqqc qacactccac catagatcac tcccctqtga ggaactactg 60
tetteacgea gaaagegtet agecatggeg tragtatgag tgtegtgeag cetecaggae 120
coccctocc gggagageca tagtggtetg eggaaceggt gagtacaceg gaattgecag 180 gacgaceggg teetttettg gateaaceg etcaatgeet ggagatttgg gegtgeeece 240
gegagaetge tageegagta gtgttgggte gegaaaggee ttgtggtaet geetgatagg 300
gtgettgega gtgeeeggg aggtetegta gaeegtgeae catgageaeg aateetaaae 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc teeggeeget tgggtggaga ggetattegg ctatgactgg gcacaacaga 480
caateggetg etetgatgee geegtgttee ggetgteage geaggggege eeggttettt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg getgetattg ggegaagtge eggggeagga teteetgtea teteacettg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tettgtcgat caggatgate tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactgtt cgccaggctc aaggcgcgca tgcccgacgg cgaggatctc gtcgtgaccc 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080 ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa eggttteeet etagegggat caatteegee eeteteeete eeeeeeeet 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320 tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgageatte ctaggggtet tteceetete geeaaaggaa tgeaaggtet gttgaatgte 1440
gtgaaggaag cagificetet ggaagettet tgaagacaaa caaegtetgt agegaceett 1500
tgcaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcggtgg agtttacctg 1920
ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980 caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040 gctcagcccg ggtacccctg gcccctctat ggcaatgagg gcttggggtg ggcaggatgg 2100 ctcctgtcac cccgtggctc tcggcctagt tggggccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggtcat cgataccete acgtgcggct tcgccgatct catggggtac 2220
atteceteg teggegeece ectagggge getgeeaggg ceetggegea tggegteegg 2280 gttetggagg acggegtgaa etatgeaaca gggaatetge eeggttgete etttetate 2340
tteettttgg etttgetgte etgtttgace ateccageit eegettatga agtgegeaac 2400
gtateeggag tgtaceatgt caegaaegae tgeteeaaeg caageattgt gtatgaggea 2460
geggacatga teatgeatae eccegggige gigeeetgeg tiegggagaa caacteetee 2520
cgctgctggg tagcgctcac tcccacgctc gcggccagga acgctagcgt ccccactacg 2580
acgatacgae gecatgtega tttgetegtt ggggeggetg etetetgete egetatgtae 2640
gtgggagate tetgeggate tgtttteete gtegeceage tgtteacett etegeetege 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760
```

```
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgcag 2820 ttactccgga tcccacaagc tgtcgtggat atggtggcgg gggcccattg gggagtccta 2880
 gegggeetig cetaetatic catggtgggg aactgggeta aggttetgat tgtgatgeta 2940
 ctctttgccg gcgttgacgg gggaacctat gtgacagggg ggacgatggc caaaaacacc 3000
 ctegggatta egtecetett tteacceggg teateceaga aaateeaget tgtaaacace 3060 aacggeaget ggsacateaa caggactgee etgaactgea atgaeteeet caacactggg 3120
 ttoottgotg ogotgttota ogtgoacaag ttoaactoat otggatgood agagogoatg 3180
gccagctgca gccccatcga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240 tcacacagct cggaccagag gccttattgt tggcactacg caccccggcc gtgcggtatc 3300 gtacccggcg cgcaggtgtg tggtccagtg tactgcttca ccccaagccc tgtcgtggtg 3360
 gggacgaccg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
 cigcttetta acaacaegeg geegeegeaa ggeaactggt tiggetgtae atggatgaat 3480
 agcactgggt tcaccaagac gtgcgggggc cccccgtgta acatcggggg gatcggcaat 3540
 aaaacettga cetgeeccae ggactgette eggaageace eegaggeeae ttacaceaag 3600
 tgtggttcgg ggccttggtt gacacccaga tgcttggtcc actacccata caggctttgg 3660
cactacecet geactgicaa etttaceate ticaaggita ggatgiaegt ggggggagtg 3720
gagcacagge tegaageege atgcaattgg actegaggag agegttgtaa cetggaggae 3780 agggacagat cagagettag eeegetgetg etgtetacaa eggagtggea ggtattgeee 3840
 tgttoottoa ccaccetace ggetetgtee actggtttga tecateteca teagaacgte 3900
gtggacgtac aatacctgta cggtataggg tcggcggttg tctcctttgc aatcaaatgg 3960
gagtatgtee tgttgetett cettettetg geggaegege gegtetgtge etgettgtgg 4020 atgatgetge tgatagetea agetgaggee geectagaga acetggtggt cetcaaegeg 4080
geateegtigg eegggegea tiggeattete teetteeteg tigttettetig tigetigeetigg 4140
tacatcaagg geaggetggt ceetggggeg geatatgeec tetaeggegt atggeegeta 4200 etcetgetee tgetggegtt accaecacga geataegeea tggaceggga gatggeagea 4260
tegtgeggag gegeggtitt egtaggtetg atactettga cettgteace geactataag 4320
etgtteeteg etaggeteat atggtggtta caatatttta teaceaggge egaggeaeae 4380
ttgcaagtgt ggatccccc cetcaacgtt egggggggce gegatgccgt catcctccte 4440
acgtgcgcga tccacccaga gctaatcttt accatcacca aaatcttgct cgccatactc 4500
ggtccactca tggtgctcca ggctggtata accaaagtgc cgtacttcgt gcgcgcacac 4560
gggctcattc gtgcatgcat gctggtgcgg aaggttgctg ggggtcatta tgtccaaatg 4620
gctetcatga agttggcege actgacaggt acgtacgttt atgaccatet caccecactg 4680
egggaetggg cecaegeggg cetaegagae ettgeggtgg eagttgagee egtegtette 4740
totgatatgg agaccaaggt tatcacctgg ggggcagaca ccgcggcgtg tggggacatc 4800
atottgggoo tgooogtoto ogooogoagg gggagggaga tacatotggg acoggcagao 4860
agectigaag ggeaggggtg gegacteete gegeetatta eggeetaete ecaacagaeg 4920 egaggeetae tiggetgeat cateactage etcacaggee gggacaggaa ecaggiegag 4980
ggggaggtcc aagtggtctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
gtgtgttgga ctgtctatca tggtgccggc tcaaagaccc ttgccggccc aaagggccca 5100
atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gcccccggg 5160 gcgcgttcct tgacaccatg cacctgcggc agctcggacc tttacttggt cacgaggcat 5220
gccgatgtca ttccggtgcg ccggcgggc gacagcaggg ggagcctact ctcccccagg 5280
ecceptetect acttgaaggg etetteggge ggteeactge tetgeeecte ggggeaeget 5340
gtgggcatct ttcgggctgc cgtgtgcacc cgaggggttg cgaaggcggt ggactttgta 5400 cccgtcgagt ctatggaaac cactatgcgg tccccggtct tcacggacaa ctcgtcccct 5460
coggocgtac cgcagacatt ccaggtggcc catctacacg cccctactgg tagoggcaag 5520
agcactaagg tgccggctgc gtatgcagcc'caagggtata aggtgcttgt'cctgaacccg 5580
teegtegeeg ceacectagg ttteggggeg tatatgteta aggeacatgg tategaeect 5640
aacatcagaa ccggggtaag gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700 ggcaagtttc ttgccgacgg tggttgctct ggggggcgct atgacatcat aatatgtgat 5760
gagtgccact caactgacte gaccactate etgggcateg gcacagtect ggaccaageg 5820
gagacggetg gagegegact egtegtgete gecaeegeta egeeteeggg ateggteaee 5880
gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
ggcaaagcca tececatega gaccateaag ggggggagge aceteatttt etgecattee 6000
aagaagaaat gtgatgagct cgccgcgaag ctgtccggcc tcggactcaa tgctgtagca 6060
tattaceggg geettgatgt atcegteata ceaactageg gagaegteat tgtegtagea 6120 aeggaegete taatgaeggg etttacegge gatttegaet eagtgatega etgeaataca 6180 tgtgteacee agaeagtega etteageetg gaeeegaeet teaecattga gaegaegaee 6240
gtgccacaag acgcggtgtc acgctcgcag cggcgaggca ggactggtag gggcaggatg 6300
ggcatttaca ggtttgtgac tecaggagaa eggceetegg gcatgttega tteeteggtt 6360 etgtgegagt getatgaege gggetgtget tggtaegage teaegeeege egagaeetea 6420
gttaggttgc gggcttacct aaacacacca gggttgcccg tctgccagga ccatctggag 6480
ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttct% gtcccagact 6540
aagcaggcag gagacaactt cccctacctg gtagcatacc aggctacggt gtgcgccagg 6600 gctcaggctc cacctccatc gtgggaccaa atgtggaagt gtctcatacg gctaaagcct 6660
acgetgeaeg ggecaaegee eetgetgtat aggetgggag eegtteaaaa egaggttaet 6720
accacacacc ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
```

```
acqaqcacct qqqtqctqqt aqqcqqaqtc ctaqcaqctc tqqccqcqta ttqcctqaca 6840
acaggcageg togtcattot gogcaggate atettoteco gaaageegge cateattese 6900
gacagggaag teetttaceg ggagttegat gagatggaag agtgegeete acaceteeet 6960
tacatogaac agggaatgca gotogoogaa caattoaaac agaaggcaat ogggttgotg 7020
caaacagcca ccaagcaagc ggaggctgct gctcccgtgg tggaatccaa gtggcggacc
ctcgaagcct tctgggcgaa gcatatgtgg aatttcatca gcgggataca atatttagca 7140
ggettgteca etetgeetgg caaceeegeg atageateac tgatggeatt cacageetet 7200
atcaccagee egeteaceae ecaacatace etectettta acateetggg gggatgggtg 7260
geogeocaac tigeteetee cagegoiget teigetiteg taggegoogg categoigga 7320
geggetgttg geageatagg eettgggaag gtgettgtgg atattttgge aggttatgga 7380
quaggggtgg caggugget cgtggcettt aaggtcatga geggegagat geeetecace 7440
gaggacetgg ttaacctact ecetgetate eteteceetg gegeeetagt egteggggte 7500
gtgtgcgcag cgatactgcg tcggcacgtg ggcccagggg agggggctgt gcagtggatg 7560 aaccggctga tagcgttcgc ttcgcggggt aaccacgtct ccccacgca ctatgtgcct 7620
gagagegaeg etgeageaeg tgteaeteag atecteteta gtettaceat cacteagetg 7680
ctgaagaggc ttcaccagtg gatcaacgag gactgctcca cgccatgctc cggctcgtgg 7740
ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggctccag 7800 tccaagctcc tgccgcgatt gccgggagtc cccttcttct catgtcaacg tgggtacaag 7860
ggagtctggc ggggcgacgg catcatgcaa accacctgcc catgtggagc acagatcacc 7920
ggacatgtga aaaacggttc catgaggatc gtggggccta ggacctgtag taacacgtgg 7980
catggaacat tesecattaa egegtacaee aegegeeeet geaegeeete eeeggegeea 8040
aattatteta gggegetgtg gegggtgget getgaggagt acgtggaggt taegegggtg 8100 ggggatttee actaegtgae gggeatgaee actgaeaacg taaagtgeee gtgteaggtt 8160
ccggcccccg aattettcac agaagtggat ggggtgcggt tgcacaggta cgctccagcg 8220
tgcaaacccc tectaeggga ggaggteaca tteetggteg ggeteaatea atacetggtt 8280
gggtcacage teccatgega geeegaaceg gaegtageag tgetcaette catgetcaee 8340 gaeeeettee acattaegge ggagaegget aagegtagge tggecagggg ateteceee 8400
feettggeea geteateage tagecagetg tetgegeett cettgaagge aacatgeact 8460
acceptcatg actocoogga egotgacoto atogaggoca acotoctgtg geggcaggag 8520
atgggcggga acatcacccg cgtggagtca gaaaataagg tagtaatttt ggactctttc 8580
gageegetee aageggagga ggatgagagg gaagtateeg tteeggegga gateetgegg 8640
aggtecagga aatteceteg agegatgeee atatgggeae geeeggatta caacceteea 8700
ctgttagagt cctggaagga cccggactac gtccctccag tggtacacgg gtgtccattg 8760 ccgcctgcca aggcccctcc gataccacct ccacggagga agaggacggt tgtcctgtca 8820
gaatctaccg tgtcttctgc cttggcggag ctcgccacaa agaccttcgg cagctccgaa 8880
tegteggeeg tegacagegg caeggeaaeg geeteteetg accageeete egacgaegge 8940
gacgcgggat ccgacgttga gtcgtactcc tccatgcccc cccttgaggg ggagccgggg 9000
gatcccgatc tcagcgacgg gtcttggtct accgtaagcg aggaggctag tgaggacgtc 9060
gtetgetget egatgteeta cacatggaca ggegeeetga teaegeeatg egetgeggag 9120
gaaaccaagc tgcccatcaa tgcactgagc aactctttgc tccgtcacca caacttggtc 9180
tatgctacaa catctcgcag cgcaagcctg cggcagaaga aggtcacctt tgacagactg 9240
caggtcctgg acgaccacta ccgggacgtg ctcaaggaga tgaaggcgaa ggcgtccaca 9300
gttaaggeta aacttetate egtggaggaa geetgtaage tgaegeeece acatteggee 9360
agatetaaat ttggetatgg ggeaaaggae gteeggaace tateeageaa ggeegttaae 9420
cacatcogot cogigiggaa ggactigoig gaagacacig agacaccaat igacaccaco 9480
atcatggcaa aaaatgaggt tttctgcgtc caaccagaga aggggggccg caagccagct 9540
cgccttatcg tattcccaga tttgggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600
gtggtctcca ccctccctca ggccgtgatg ggctcttcat acggattcca atactctcct 9660
ggacageggg tegagtteet ggtgaatgee tggaaagega agaaatgeee tatgggette 9720
gcatatgaca cccgctgttt tgactcaacg gtcactgaga atgacatccg tgttgaggag 9780
tcaatctacc aatgttgtga cttggcccc gaagccagac aggccataag gtcgctcaca 9840 gagcggcttt acatcggggg ccccctgact aattctaaag ggcagaactg cggctatcgc 9900
eggtgeegeg egageggtgt actgaegace agetgeggta atacceteae atgttaettg 9960
aaggeegetg eggeetgteg agetgegaag etecaggaet geaegatget egtatgegga 10020
gacgaccttg tcgttatctg tgaaagcgcg gggacccaag aggacgaggc gagcctacgg 10080 gccttcacgg aggctatgac tagatactct gcccccctg gggacccgcc caaaccagaa 10140
tacgacttgg agttgataac atcatgctcc tocaatgtgt cagtcgcgca cgatgcatct 10200
ggcaaaaggg tgtactatct caccegtgac cecaccacce cecttgegeg ggctgegtgg 10260 gagacageta gacacactec agtcaattee tggctaggea acateateat gtatgegeec 10320
accttgtggg caaggatgat cetgatgact catttettet ceateettet ageteaggaa 10380
caacttgaaa aagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440
gacctacctc agatcattca acgactccat ggccttagcg cattttcact ccatagttac 10500
tctccaggtg agatcaatag ggtggcttca tgcctcagga aacttggggt.accgcccttg 10560 cgagtctgga gacatcgggc cagaagtgtc cgcgctaggc tactgtccca gggggggagg 10620
gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680
ccaatcccgg ctgcgtccca gttggattta tccagctggt tcgttgctgg ttacagcggg 10740
ggagacatat atcacageet gtetegtgee egaceeeget ggtteatgtg gtgeetaete 10800
```

#### CA 02303526 2000-03-31

ctccaggcca	taggggtagg ataggccatc ttttttttc	ctatttttt	ccctttttt	ttttctttt	tttttttt	10920
tagctccatc	ttagccctag	tcacggctag	ctgtgaaagg	tccgtgagcc	gcttgactgc	11040 11076

<210> 2

```
<211> 8637
<212> DNA
<213> Hepatitis C Virus
<400> 2
gccagcccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tetteacgea gaaagegtet agecatggeg ttagtatgag tgtegtgeag cetecaggae 120
ccccccccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg teettettg gateaacccg eteaatgeet ggagatttgg gegtgeece 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttccgc ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540
acctgtccgg tgccctgaat gaactgcagg acgaggcagc gcggctatcg tggctggcca 600
cgacgggcgt teettgegea getgtgeteg acgttgteae tgaageggga agggaetgge 660 tgetattggg cgaagtgeeg gggeaggate teetgteate teacettget cetgeegaga 720
aagtatecat catggetgat geaatgegge ggetgeatae gettgateeg getaeetgee 780
cattegacea ceaagegaaa categeateg agegageaeg taeteggatg gaageeggte 840
ttgtcgatca ggatgatctg gacgaagagc atcaggggct cgcgccagcc gaactgttcg 900 ccaggctcaa ggcgcgcatg cccgacggcg aggatctcgt cgtgacccat ggcgatgcct 960
qcttqccqaa tatcatggtg gaaaatggcc gcttttctgg attcatcgac tgtggccggc 1920
tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140 agcgcatcgc cttctatcgc cttcttgacg agttcttctg agtttaaaca gaccacaacg 1200
gttteeetet agegggatea atteegeece teteeeteee eeceecetaa egttactgge 1260
cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320 ccgtcttttg gcaatgtgag ggcccggaaa cctggccctg tcttcttgac gagcattcct 1380 aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440
gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg 1500
aaccccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560
gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1620
tggctctcct caagegtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1680
atgggatctg atctggggcc tcggtgcaca tgctttacat gtgtttagtc gaggttaaaa 1740
aacgtctagg ccccccgaac cacggggacg tggttttcct ttgaaaaaca cgataatacc 1800
atggaccggg agatggcagc atcgtgcgga ggcgcggttt tcgtaggtct gatactcttg 1860
accttgtcac cgcactataa gctgttcctc gctaggctca tatggtggtt acaatatttt 1920
atcaccaggg ccgaggcaca cttgcaagtg tggatccccc ccctcaacgt tcgggggggc 1980
cgcgatgccg tcatcctcct cacgtgcgcg atccacccag agctaatctt taccatcacc 2040
aaaatettge tegecataet eggteeacte atggtgetee aggetggtat aaccaaagtg 2100
cogtacttcg tgcgcgcaca cgggctcatt cgtgcatgca tgctggtgcg gaaggttgct 2160
gggggtcatt atgtccaaat ggctctcatg aagttggccg cactgacagg tacgtacgtt 2220
tatgaccatc tcaccccact gegggactgg gcccacgegg gcctacgaga ccttgcggtg 2280
geagttgage cegtegtett etetgatatg gagaceaagg ttateacetg gggggeagae 2340 accgeggegt gtggggaeat catettggge etgecegtet eegeeggag ggggagggag 2400
atacatctgg gaccggcaga cagccttgaa gggcaggggt ggcgactcct cgcgcctatt 2460
acggcctact cccaacagac gcgaggccta cttggctgca tcatcactag cctcacaggc 2520
cgggacagga accaggtcga ggggggggtc caagtggtct ccaccgcaac acaatctttc 2580
ctggcgacct gcgtcaatgg cgtgtgttgg actgtctatc atggtgccgg ctcaaagacc 2640 cttgccggcc caaagggccc aatcacccaa atgtacacca atgtggacca ggacctcgtc 2700
ggctggcaag cgccccccgg ggcgcgttcc ttgacaccat gcacctgcgg cagctcggac 2760
ctttacttgg tcacgaggca tgccgatgtc attccggtgc gccggcgggg cgacagcagg 2820
gggagcctac tetececcag geoegtetee tacttgaagg getetteggg eggtecactg 2880 etetgeceet eggggeaege tgtgggeate tttegggetg eegtgtgeae eegaggggtt 2940
gcgaaggcgg tggactttgt accogtogag totatggaaa coactatgcg gtocccggto 3000
ttcacggaca actcgtcccc tccggccgta ccgcagacat tccaggtggc ccatctacac 3060
gecectactg gtageggeaa gageactaag gtgeeggetg egtatgeage ceaagggtat 3120
aaggtgettg teetgaacee gteegtegee gecaceetag gtttegggge gtatatgtet 3180
aaggcacatg gtatcgaccc taacatcaga accggggtaa ggaccatcac cacgggtgcc 3240
cccatcacgt actccaccta tggcaagttt cttgccgacg gtggttgctc tgggggcgcc 3300 tatgacatca taatatgtga tgagtgccac tcaactgact cgaccactat cctgggcatc 3360
ggcacagtcc tggaccaagc ggagacggct ggagcgcgac tcgtcgtgct cgccaccgct 3420
acgecteegg gateggteae egtgeeacat ceaaacateg aggaggtgge tetgteeage 3480
actggagaaa toccotttta tggcaaagec atccccatcg agaccatcaa gggggggagg 3540 cacctcattt totgccatto caagaagaaa tgtgatgago togcogogaa gotgtooggo 3600
```

```
ctcggactca atgctgtage atattaccgg ggccttgatg tatccgtcat accaactage 3660
ggagacgtca ttgtcgtage aacggacget ctaatgacgg getttacegg egatttegae 3720 teagtgateg actgeaatae atgtgtcace cagacagteg acttcageet ggaceegace 3780
ttcaccattg agacgacgac cgtgccacaa gacgcggtgt cacgctcgca gcggcgaggc 3840
aggactggta ggggcaggat gggcatttac aggtttgtga ctccaggaga acggccctcg
ggcatgttcg attecteggt tetgtgegag tgctatgacg egggetgtge ttggtacgag 3960
ctcacgcccg ccgagacctc agttaggttg cgggcttacc taaacacacc agggttgccc 4020
gtotgocagg accatotgga gttotgggag agogtottta caggootcae coacatagae 4080
geceatttet tgteccagae taageaggea ggagaeaaet teeeetaeet ggtageatae 4140
caggetacge tgtgegecag ggeteagget ccacetecat egtgggacea aatgtggaag 4200 tgteteatae ggetaaagee taegetgeae gggeeaaege eeetgetgta taggetggga 4260
googttoaaa acgaggttac taccacacac cocataacca aatacatcat ggoatgoatg 4320
teggetgace tggaggtegt cacgageace tgggtgetgg taggeggagt cetageaget 4380
ctggccgcgt attgcctgac aacaggcagc gtggtcattg tgggcaggat catcttgtcc 4440 ggaaagccgg ccatcattcc cgacagggaa gtcctttacc gggagttcga tgagatggaa 4500
gagtgcgcct cacacctccc ttacatcgaa cagggaatgc agctcgccga acaattcaaa 4560
cagaaggcaa togggttgot gcaaacagco accaagcaag oggaggotgo tgotocogtg 4620
gtggaatcca agtggcggac cctcgaagcc ttctgggcga agcatatgtg gaatttcatc 4680
agogggatac aatatttago aggottgtoc actotgootg gcaaccoogo gatagoatca 4740
ctgatggcat tcacagecte tateaceage cegeteacea eccaacatae ecteetgttt 4800
aacateetgg ggggatgggt ggeegeecaa ettgeteete eeagegetge ttetgettte 4860
gtaggegeeg geategetgg ageggetgtt ggeageatag geettgggaa ggtgettgtg 4920 gatattttgg caggttatgg ageaggggtg geaggegege tegtggeett taaggteatg 4980
ageggegaga tgeecteeae egaggaeetg gttaacetae teeetgetat eeteteeeet 5040
ggcgccctag tcgtcgggt cgtgtgcgca gcgatactgc gtcggcacgt gggcccaggg 5100 gagggggctg tgcagtggat gaaccggctg atagcgttcg cttcgcgggg taaccacgtc 5160 tcccccacge actatgtgcc tgagagcgac gctgcagcac gtgtcactca gatcctctct 5220
agtettacca teacteaget getgaagagg etteaceagt ggateaacga ggactgetee 5280
acgccatgct coggctogtg gctaagagat gtttgggatt ggatatgcac ggtgttgact 5340
gatttcaaga cotggotoca gtocaagoto otgoogogat tgoogogagt cocottotto 5400
tcatgtcaac gtgggtacaa gggagtctgg cggggcgacg gcatcatgca aaccacctgc 5460
ccatgtggag cacagatcac cggacatgtg aaaaacggtt ccatgaggat cgtggggcct 5520
aggacctgta gtaacacgtg gcatggaaca ttccccatta acgcgtacac cacgggcccc 5580
tgcacgcct ccccggcgc aaattattct agggcgctgt ggcgggtggc tgctgaggag 5640 tacgtggagg ttacgcgggt gggggatttc cactacgtga cgggcatgac cactgacaac 5700 gtaaagtgcc cgtgtcaggt tccggcccc gaattcttca cagaagtgga tggggtgcgg 5760
ttgcacaggt acgctccagc gtgcaaaccc ctcctacggg aggaggtcac attcctggtc 5820
gggctcaatc aatacctggt tgggtcacag ctcccatgcg agcccgaacc ggacgtagca 5880
gtgctcactt ccatgctcac cgacccctcc cacattacgg cggagacggc taagcgtagg 5940
ctggccaggg gatctccccc ctccttggcc agctcatcag ctagccagct gtctgcgcct 6000
teettgaagg caacatgeae taccegteat gaeteeeegg aegetgaeet categaggee 6060
aacctcctgt ggcggcagga gatgggcggg aacatcaccc gcgtggagtc agaaaataag 6120 gtagtaattt tggactcttt cgagccgctc caagcggagg aggatgagag ggaagtatcc 6180
gttccggcgg agatcctgcg gaggtccagg aaattccctc gagcgatgcc catatgggca 6240
cgcccggatt acaaccctcc actgttagag teetggaagg acceggacta cgtccctcca 6300
gtggtacacg ggtgtccatt gccgcctgcc aaggcccctc cgataccacc tccacggagg 6360
aagaggacgg tigtcctgtc agaatctacc gtgtcttctg ccttggcgga gctcgccaca 6420 aagaccttcg gcagctccga atcgtcggcc gtcgacagcg gcacggcaac ggcctctcct 6480
gaccagecet eegacgacgg egacgeggga teegacgttg agtegtacte etecatgeee 6540
eccettgagg gggageeggg ggateeegat eteagegaeg ggtettggte tacegtaage 6600
gaggaggeta gigaggaegi egictgetge tegatgieci acacatggae aggegeeeig 6660
atcacgccat gegetgegga ggaaaccaag etgeccatca atgeactgag caactetttg 6720
ctccgtcacc acaacttggt ctatgctaca acatctcgca gcgcaagcct gcggcagaag 6780
aaggtcacct ttgacagact geaggteetg gacgaccact accgggacgt getcaaggag 6840
atgaaggega aggegtécae agttaagget aaaettetat eegtggagga ageetgtaag 6900
ctgacgcccc cacattegge cagatetaaa tttggetatg gggeaaagga egteeggaae 6960
ctatecagea aggeogttaa ecacateege teegtgtgga aggaettget ggaagacaet 7020
gagacaccaa ttgacaccac catcatggca aaaaatgagg ttttctgcgt ccaaccagag 7080
aaggggggcc gcaagccagc tcgccttatc gtattcccag atttgggggt tcgtgtgtgc 7140 gagaaaatgg ccctttacga tgtggtctcc accetecete aggccgtgat gggctcttca 7200
tacggattcc aatactetee tggacagegg gtegagttee tggtgaatge etggaaageg 7260
aagaaatgcc ctatgggctt cgcatatgac acccgctgtt ttgactcaac ggtcactgag 7320
aatgacatee gtgttgagga gtcaatetae caatgttgtg acttggeece egaageeaga 7380
caggecataa ggtegeteae agageggett tacategggg geeeeetgae taattetaaa 7440
gggcagaact gcggctatcg ccggtgccgc gcgagcggtg tactgacgac cagctgcggt 7500
aataccetca catgitacti gaaggeeget geggeetgte gagetgegaa geteeaggae 7560 tgeacgatge tegtatgegg agacgacett giegttatet gigaaagege ggggaeecaa 7620
```

паппаспаци	cgagcctacg	ggccttcacg	gaggctatga	ctagatactc	tgcccccct	7680
ggggacccgc	ccaaaccaga	atacgacttg	gagttgataa	catcatgctc	ctccaatgtg	//40
tcagtcgcgc	acgatgcatc	tggcaaaagg	gtgtactatc	tcacccgtga	ccccaccacc	7800
cccttqcqc	gggctgcgtg	ggagacagct	agacacactc	cagtcaattc	ctggctaggc	7860
aacatcatca	tgtatgcgcc	caccttgtgg	gcaaggatga	tcctgatgac	tcatttcttc	7920
tccatccttc	tagctcagga	acaacttgaa	aaagccctag	attgtcagat	ctacggggcc	7980
tottactcca	ttgagccact	tgacctacct	cagatcattc	aacgactcca	tggccttagc	8040
gcattttcac	tccatagtta	ctctccaggt	gagatcaata	gggtggcttc	atgcctcagg	8100
aaacttgggg	taccgccctt	gcgagtctgg	agacatcggg	ccagaagtgt	ccgcgctagg	8160
ctactgtccc	aggggggag	ggctgccact	tgtggcaagt	acctcttcaa	ctgggcagta	8220
aggaccaage.	tcaaactcac	tecaateceg.	gctgcgtccc	agttggattt	atccagctgg	8280
ttcattacta	gttacagcgg	gggagacata	tatcacagcc	tgtctcgtgc	ccgaccccgc	8340
togttcatgt	ggtgcctact	cctactttct	gtaggggtag	gcatctatct	actccccaac	8400
cgatgaacgg	ggagctaaac	actccaggcc	aataggccat	cctgtttttt	tcccttttt	8460
tttttcttt	tttttttt	ttttttttt	tttttttt	ctcctttttt	tttcctcttt	8520
ttttccttt	ctttcctttg	gtggctccat	cttagcccta	gtcacggcta	gctgtgaaag	8580
gtccgtgagc	cgcttgactg	cagagagtgc	tgatactggc	ctctctgcag	atcaagt	8637

```
<211> 8001
<212> DNA
<213> Hepatitis C Virus
<400> 3
gecagecece gattggggge gacactecae catagateae teceetgtga ggaactaetg 60
tetteaegea gaaagegtet ageeatggeg ttagtatgag tgtegtgeag eeteeaggae 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg teetttettg gateaacccg eteaatgeet ggagatttgg gegtgeecce 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caatoggotg stotgatgos googtgttos ggotgtcage gcaggggege coggttottt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
egtggetgge caegaeggge gtteettgeg eagetgtget egaegttgte actgaagegg 660
gaagggactg getgetattg ggegaagtge eggggeagga teteetgtea teteacettg 720
ctectgeega gaaagtatee ateatggetg atgeaatgeg geggetgeat acgettgate 780 eggetacetg eccattegae caccaagega aacategeat egagegagea egtactegga 840
tggaagccgg tottgtcgat caggatgate tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactgtt cgccaggctc aaggcgcgca tgcccgacgg cgaggatete gtcgtgacce 960
atggcgatge etgettgeeg aatateatgg tggaaaatgg eegettttet ggatteateg 1020 aetgtggeeg getgggtgt geggaeeget ateaggaeat agegttgget accegtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tecaccatat tgeogtottt tggcaatgtg agggooogga aacctggooc tgtottottg 1380
acgageatte ctaggggtet tteceetete gecaaaggaa tgeaaggtet gttgaatgte 1440
gtgaaggaag cagtteetet ggaagettet tgaagacaaa caacgtetgt agegaceett 1500 tgeaggeage ggaacceec acetggegae aggtgeetet geggeeaaaa geeaegtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatt gtatgggate tgatetgggg ceteggtgea catgetttae atgtgtttag 1740
togaggttaa aaaacgtota ggoocoocga accaegggga egtggtttto otttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcatcatca ctagcctcac aggccgggac aggaaccagg tcgaggggga ggtccaagtg 1920
gtetecaceg caacacaate titectggeg acetgegtea atggegtgtg tiggactgte 1980 tateatggtg eeggeteaaa gaceettgee ggeecaaagg geecaateae eeaaatgtae 2040
accaatgigg accaggacet egicggeigg caagegeeee eeggggegeg tieettigaea 2100
ccatgcacct gcggcagete ggacetttae ttggtcacga ggcatgccga tgtcattecg 2160
gtgcgccggc ggggcgacag cagggggagc ctactctccc ccaggcccgt ctcctacttg 2220
aagggetett egggeggtee aetgetetge eeetegggge aegetgtggg catetttegg 2280
getgeegtgt geaceegagg ggttgegaag geggtggaet ttgtaceegt egagtetatg 2340
gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400 acattccagg tggcccatct acacgccct actggtagcg gcaagagcac taaggtgccg 2460 gctgcgtatg cagcccaagg gtataaggtg cttgtcctga accegtccgt cgccgccacc 2520
ctaggtttcg gggcgtatat gtctaaggca catggtatcg accetaacat cagaaccggg 2580
gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtggtt gctctggggg cgcctatgac atcataatat gtgatgagtg ccactcaact 2700
gactogacca ctatootggg catoggoaca gtootggaco aagoggagac ggotggagog 2760
cgactogtog tgctcgccac cgctacgcct ccgggatcgg tcaccgtgcc acatccaaac 2820
ategaggagg tggetetgte eageaetgga gaaateeeet tttatggeaa ageeateeee 2880
atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940 gagctcgccg cgaagctgtc cggcctcgga ctcaatgctg tagcatatta ccggggcctt 3000
gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaatg 3060
acgggcttta ccggcgattt cgactcagtg atcgactgca atacatgtgt cacccagaca 3120
gtegaettea geetggaeee gaeetteaee attgagaega egaeegtgee acaagaegeg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtagggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctcgggcatg ttcgattcct cggttctgtg cgagtgctat 3300
gacgeggget gtgettggta egageteaeg eeegeegaga eeteagttag gttgeggget 3360
tacctaaaca caccagggtt gcccgtctgc caggaccatc tggagttctg ggagagcgtc 3420 tttacaggcc tcacccacat agacgcccat ttcttgtccc agactaagca ggcaggagac 3480
aacttcccct acctggtage ataccagget acggtgtgcg ccagggetca ggctccacct 3540
```

```
ccatcgtggg accaaatgtg gaagtgtete atacggetaa ageetaeget geaegggeea 3600
acgecectge tgtatagget gggageegtt caaaacgagg ttactaceae acaceccata 3660
accaaataca tcatggcatg catgtoggct gacctggagg tcgtcacgag cacctgggtg 3720
ctggtaggcg gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggtc 3780
attgtgggca ggatcatctt gtccggaaag ccggccatca ttcccgacag ggaagtcctt 3840 taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900 atgcagctcg ccgaacaatt caaacagaag gcaatcgggt tgctgcaaac agccaccaag 3960
caageggagg etgetgetee egtggtggaa tecaagtgge ggaceetega ageettetgg 4020 gegaageata tgtggaattt cateageggg atacaatatt tageaggett gtecaetetg 4080
cetggcaace cegegatage atcactgatg gcatteacag cetetateae cageeegete 4140 accacecaac ataceteet gtttaacate etggggggat gggtggeege ceaacttget 4200
cotoccagog otgottotgo titogtaggo geoggoateg otggageggo tgttggcage 4260
ataggeettg ggaaggtget tgtggatatt ttggcaggtt atggageagg ggtggeagge 4320
gegetegtgg cetttaaggt catgagegge gagatgeet ceacegagga cetggttaac 4380 etactecetg etateetete eeetggegee etagtegteg gggtegtgtg egeagegata 4440
ctgcgtcggc acgtgggccc aggggaggg gctgtgcagt ggatgaaccg gctgatagcg 4500
ttegettege ggggtaacea egteteeee acgeaetatg tgeetgagag egaegetgea 4560 geaegtgtea eteagateet etetagtett accateacte agetgetgaa gaggetteae 4620
cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680
gattggatat gcacggtgtt gactgatttc aagacctggc tccagtccaa gctcctgccg 4740
cgattgccgg gagtcccctt cttctcatgt caacgtgggt acaagggagt ctggcggggc 4800
gacggcatca tgcaaaccac ctgcccatgt ggagcacaga tcaccggaca tgtgaaaaac 4860 ggttccatga ggatcgtggg gcctaggacc tgtagtaaca cgtggcatgg aacattcccc 4920
attaacgcgt acaccacggg cccctgcacg ccctccccgg cgccaaatta ttctagggcg 4980 ctgtggcggg tggctgctga ggagtacgtg gaggttacgc gggtggggga tttccactac 5040 gtgacggca tgaccactga caacgtaaag tgcccgtgte aggttccggc ccccgaattc 5100 ttcacagaag tggatgggg gcggttgcac aggtacgct cagcgtgcaa acccctccta 5160
cgggaggagg tcacattcct ggtcgggctc aatcaatacc tggttgggtc acagctccca 5220
tgcgagcccg aaccggacgt agcagtgctc acttccatgc tcaccgaccc ctcccacatt 5280
acggeggaga eggetaageg taggetggee aggggatete ecceeteett ggeeagetea 5340
tragctager agetgeetge gentlecting aaggraarat gractarreg tratgarter 5400
ccggacgctg acctcatcga ggccaacctc ctgtggcggc aggagatggg cgggaacatc 5460
accegegteg agteagaaa taaggtagta attitiggaet etttegagee geteeaageg 5520
gaggaggatg agagggaagt atcogttocg goggagatoc tgoggaggto caggaaatto 5580
cctcgagcga tgcccatatg ggcacgcccg gattacaacc ctccactgtt agagtcctgg 5640
aaggaccegg actacgteec tecagtggta caegggtgte cattgeegee tgeeaaggee 5700
cotécgatac cacetécacy gaggaagagg acggttgtec tyteagaate tacegtgtet 5760 tetgeettgg eggagetege cacaaagace tteggeaget ecgaategte ggeegtegae 5820
ageggeaegg caaeggeete teetgaeeag eesteegaeg aeggegaege gggateegae 5880
gttgagtcgt actoctocat gcccccctt gagggggagc cgggggatcc cgatctcagc 5940
gacgggtctt ggtctaccgt aagcgaggag gctagtgagg acgtcgtctg ctgctcgatg 6000
tectacaeat ggacaggege cetgateaeg ceatgegetg eggaggaaae caagetgeee 6060
atcaatgcac tgagcaacte titgeteegt caccacaact tggtetatge tacaacatet 6120
cgcagcgcaa gcctgcggca gaagaaggtc acctttgaca gactgcaggt cctggacgac 6180
cactaccggg acgtgctcaa ggagatgaag gcgaaggcgt ccacagttaa ggctaaactt 6240
ctatccgtgg aggaagcctg taagctgacg ccccacatt cggccagatc taaatttggc 6300 tatggggcaa aggacgtccg gaacctatcc agcaaggccg ttaaccacat ccgctccgtg 6360
tggaaggact tgctggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
gaggttttct gcgtccaacc agagaagggg ggccgcaagc cagctcgcct tatcgtattc 6480
ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggt ctccaccctc 6540 cctcaggccg tgatgggctc ttcatacgga ttccaatact ctcctggaca gcgggtcgag 6600
ttcctggtga atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacacccgc 6660
tgttttgact caacggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt 6720
tgtgacttgg cccccgaagc cagacaggcc ataaggtcgc tcacagagcg gctttacatc 6780
gggggcccc tgactaattc taaagggcag aactgcggct atcgccggtg ccgcgcgagc 6840
ggtgtactga cgaccagctg cggtaatacc ctcacatgtt acttgaaggc cgctgcggcc 6900
tgtcgagctg cgaagctcca ggactgcacg atgctcgtat gcggagacga ccttgtcgtt 6960
atotgtgaaa gogogggac ccaagaggac gaggogagec tacgggcett cacggagget 7020 atgactagat actotgoccc coctggggac cogcocaaac cagaatacga cttggagttg 7080
ataacatcat gctcctccaa tgtgtcagtc gcgcacgatg catctggcaa aagggtgtac 7140
tateteacee gtgaceeeae caceeeett gegegggetg egtgggagae agetagaeae 7200
actocagtoa attoctggot aggoaacato atcatgtatg coccacett gigggoaagg 7260
atgatectga tgaeteattt etteteeate ettetagete aggaacaaet tgaaaaagee 7320
ctagattgtc agatctacgg ggcctgttac tccattgagc cacttgacct acctcagatc 7380
attcaacgae tecatggeet tagegeattt teactceata gttaetetee aggtgagate 7440
aatagggtgg cttcatgcct caggaaactt ggggtaccgc ccttgcgagt ctggagacat 7500 cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560
```

aagtacctct	tcaactgggc	agtaaggacc	aagctcaaac	tcactccaat	cccggctgcg	7620
tcccagttgg	atttatccag	ctggttcgtt.	gctggttaca	gcgggggaga	catatatcac	7680
					ttctgtaggg	
gtaggcatct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctgtt	tttttccctt	ttttttttc	tttttttt	tttttttt	tttttttt	7860
ttttctcctt	ttttttcct	cttttttcc	ttttctttcc	tttggtggct	ccatcttagc	7920
cctagtcacg	gctagctgtg	aaaggtccgt	gagecgettg	actgcagaga	gtgctgatac	7980
tagcctctct	gcagatcaag	t		•		8001

```
<210> 4
<211> 7989
<212> DNA
<213> Hepatitis C Virus
```

```
<400> 4
gecagecece gattggggge gaeactecae catagateae teceetgtga ggaactaetg 60
tettcacgca gaaagegtet agecatggeg ttagtatgag tgtcgtgcag cetecaggae 120 eccectece gggagageca tagtggtetg eggaaceggt gagtacaceg gaattgecag 180
gacgaccggg tecttettg gateaacccg eteaatgeet ggagatttgg gegtgeecce 240
gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
gtgettgega gtgeeceggg aggtetegta gaeegtgeae catgageaeg aateetaaae 360
ctcaaaqaaa aaccaaaggg cgcgccatga ttgaacaaga tggattgcac gcaggttctc 420
cggccgcttg ggtggagagg ctattcggct atgactgggc acaacagaca atcggctgct 480
ctgatgccgc cgtgttccgg ctgtcagcgc aggggcgccc ggttcttttt gtcaagaccg 540
acctgiccgg tgecetgaat gaactgeagg acgaggeage geggetateg tggetggeea 600 egaegggegt teettgegea getgtgeteg acgttgteae tgaageggga agggaetgge 660
tgctattggg cgaagtgccg gggcaggatc tcctgtcatc tcaccttgct cctgccgaga 720
aagtatocat catggotgat goaatgoggo ggotgoatao gottgatoog gotacotgoo 780
cattegacea ceaagegaaa categeateg agegageaeg tacteggatg gaageeggte 840 ttgtegatea ggatgatetg gacgaagage ateagggget egegeeagee gaaetgtteg 900
ccaggeteaa ggegegeatg ceegaeggeg aggatetegt egtgaeceat ggegatgeet 960
gettgeegaa tateatggtg gaaaatggee gettttetgg atteategae tgtggeegge 1020
tgggtgtggc ggaccgctat caggacatag cgttggctac ccgtgatatt gctgaagagc 1080 ttggcggcga atgggctgac cgcttcctcg tgctttacgg tatcgccgct cccgattcgc 1140
agegeatege ettetatege ettettgaeg agttettetg agtttaaaca gaecacaaeg 1200
gtttccctct agcgggatca attccgcccc tctccctccc ccccccctaa cgttactggc 1260
cgaagccgct tggaataagg ccggtgtgcg tttgtctata tgttattttc caccatattg 1320
cogtettitg geaatgtgag ggeeeggaaa eetggeeetg tettettgae gageatteet 1380
aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca 1440
gtteetetgg aagettettg aagacaaaca aegtetgtag egaceetttg caggcagegg 1500
aacccccac ctggcgacag gtgcctctgc ggccaaaagc cacgtgtata agatacacct 1560
gcaaaggcgg cacaacccca gtgccacgtt gtgagttgga tagttgtgga aagagtcaaa 1620
tggctctcct caagcgtatt caacaagggg ctgaaggatg cccagaaggt accccattgt 1680
atgggatetg atctggggee teggtgeaca tgetttaeat gtgtttagte gaggttaaaa 1740 aacgtetagg eecceegaac caeggggaeg tggtttteet ttgaaaaaca egataataee 1800
atggcgccta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 1860
agecteacag geegggacag gaaccaggte gagggggagg tecaagtggt etecacegea 1920
acacaatett teetggegae etgegteaat ggegtgtgtt ggaetgteta teatggtgee 1980
ggotcaaaga coottgoogg occaaagggo ocaatoacco aaatgtacac caatgtggac 2040
caggaceteg teggetggea agegeeece ggggegegtt cettgacace atgeacetge 2100
ggcagctcgg acctttactt ggtcacgagg catgccgatg tcattccggt gcgccggcgg 2160
ggcgacagca gggggagcct actotococc aggcccgtct cotacttgaa gggctcttcg 2220
ggcggtccac tgctctgccc ctcggggcac gctgtgggca tctttcgggc tgccgtgtgc 2280
accegagggg ttgcgaaggc ggtggacttt gtaccegtcg agtctatgga aaccactatg 2340 cggtccccgg tcttcacgga caactcgtcc cctccggccg taccgcagac attccaggtg 2400
gcccatctac acgcccctac tggtagcggc aagagcacta aggtgccggc tgcgtatgca 2460
geceaagggt ataaggtget tgteetgaac eegteegteg eegecaeeet aggttteggg 2520 gegtatatgt etaaggeaca tggtategae eetaacatea gaaceggggt aaggaeeate 2580
accaegggtg cececateae gtaetecaee tatggeaagt ttettgeega eggtggttge 2640
tctgggggcg cctatgacat cataatatgt gatgagtgcc actcaactga ctcgaccact 2700
atectgggea teggeacagt eetggaecaa geggagaegg etggagegeg aetegtegtg 2760
ctogocacog ctacgoctoc gggatoggto acogtgocac atocaaacat cgaggaggtg 2820
gctctgtcca gcactggaga aatccccttt tatggcaaag ccatccccat cgagaccatc 2880
aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgcg 2940
aagetgteeg geeteggaet caatgetgta geatattaee ggggeettga tgtateegte 3000
ataccaacta geggagaegt cattgtegta geaacggaeg etetaatgae gggetttaee 3060
ggcgatttcg actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcagc 3120
ctggaccega cetteaceat tgagaegaeg acegtgeeae aagaegeggt gteaegeteg 3180
cagoggogag geaggaetgg taggggeagg atgggeattt acaggtttgt gaetecagga 3240 gaacggeett egggeatgtt egatteeteg gttetgtgeg agtgetatga egegggetgt 3300
gettggtacg agetcacgee egeegagace teagttaggt tgegggetta cetaaacaca 3360
ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagcgtctt tacaggcctc 3420
acceacatag aegeceattt ettgteecag actaageagg caggagacaa etteecetae 3480 etggtageat accaggetae ggtgtgegee agggeteagg etceacetee ategtgggae 3540
caaatgigga agtgictcat acggctaaag cotacgctgc acgggccaac gecectgctg 3600
```

```
tataggetgg gageegttea aaaegaggtt actaceaeae acceeataae caaatacate 3660
atggcatgca tgtcggctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 3720
gtoctagoag etetggooge grattgootg acaacaggoa gegtggteat tgtgggoagg 3780
atcatcttgt coggaaagee ggecateatt coogacaggg aagteettta cogggagtte 3840
qatqaqatqq aaqaqtqcqc ctcacacctc ccttacatcq aacaqqqaat gcagctcgcc 3900
gaacaattca aacagaagge aategggttg etgeaaacag ecaceaagea ageggagget 3960
getgeteceg tggtggaate caagtggegg accetegaag cettetggge gaageatatg 4020
tggaatttca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4080
gegatageat cactgatgge atteacagee tetateacea gecegeteae cacceaacat 4140
accetectgt ttaacatect ggggggatgg gtggeegeec aacttgetee teceageget 4200 gettetgett tegtaggege eggeateget ggageggetg ttggeageat aggeettggg 4260
aaggtgcttg tggatatttt ggcaggttat ggagcagggg tggcaggcgc gctcgtggcc 4320 .
tttaaggtca tgagcggcga gatgccctcc accgaggacc tggttaacct actccctgct 4380
atcetetece etggegeet agtegteggg gtegtgtgeg cagegatact gegteggeac 4440
gtgggcccag gggaggggc tgtgcagtgg atgaaccggc tgatagcgtt cgcttcgcgg 4500
ggtaaccacg tetececeae geactatgtg cetgagageg acgetgeage acgtgteact 4560
cagatectet etagtettae cateacteag etgetgaaga ggetteacea gtggateaae 4620
gaggactget ccaegecatg etceggeteg tggetaagag atgtttggga ttggatatge 4680 aeggtgttga etgattteaa gacetggete eagteeaage teetgeegeg attgeeggga 4740
gteeeettet teteatgtea acgtgggtae aagggagtet ggeggggega eggeateatg 4800
caaaccacct gcccatgtgg agcacagatc accggacatg tgaaaaacgg ttccatgagg 4860
atcgtggggc ctaggacctg tagtaacacg tggcatggaa cattccccat taacgcgtac 4920
accaegggee cetgeacgee eteceeggeg ceaaattatt etagggeget gtggegggtg 4980
gctgctgagg agtacgtgga ggttacgcgg gtgggggatt tccactacgt gacgggcatg 5040
accactgaca acgtaaagtg cocgtgtcag gttccggccc ccgaattett cacagaagtg 5100
gatggggtgc ggttgcacag gtacgctcca gcgtgcaaac ccctcctacg ggaggaggtc 5160 acattcctgg tcgggctcaa tcaatacctg gttgggtcac agetcccatg cgagcccgaa 5220
coggacgtag cagtgotoac ttocatgoto accgaccoot cocacattac ggcggagacg 5280
gctaagcgta ggctggccag gggatctccc ccctccttgg ccagctcatc agctagccag 5340
ctgtctgcgc cttccttgaa ggcaacatgc actacccgtc atgactcccc ggacgctgac 5400 ctcatcgagg ccaacctcct gtggcggcag gagatgggcg ggaacatcac ccgcgtggag 5460 tcagaaaata aggtagtaat tttggactct ttcgagccgc tccaagcgga ggaggatgag 5520
agggaagtat ccgttccggc ggagatcctg cggaggtcca ggaaattccc tcgagcgatg 5580 cccatatgg cacgcccgga ttacaaccct ccactgttag agtcctggaa ggacccggac 5640
tacgtccctc cagtggtaca cgggtgtcca ttgccgcctg ccaaggcccc tccgatacca 5700
cctccacgga ggaagaggac ggttgtcctg tcagaatcta ccgtgtcttc tgccttggcg 5760
gagetegeca caaagacett eggeagetee gaategtegg eegtegacag eggeaeggea 5820
acggeetete etgaceagee etcegaegae ggegaegegg gateegaegt tgagtegtae 5880 teeteeatge ecceeettga gggggageeg ggggateeeg ateteagega egggtettgg 5940
totacogtaa gogaggagge tagtgaggac gtogtotget gotogatgto ctacacatgg 6000
acaggegeee tgateaegee atgegetgeg gaggaaacea agetgeeeat caatgeaetg 6060
ageaactett tgeteegtea ecacaacttg gtetatgeta caacateteg cagegeaage 6120
ctgcggcaga agaaggtcac ctttgacaga ctgcaggtcc tggacgacca ctaccgggac 6180
gtgctcaagg agatgaaggc gaaggcgtcc acagttaagg ctaaacttct atccgtggag 6240
gaageetgta agetgaegee eccaeatteg geeagateta aatttggeta tggggcaaag 6300
gacgteegga acctateeag caaggeegtt aaccaeatee geteegtgtg gaaggaettg 6360 etggaagaca etgagaeace aattgaeace accateatgg caaaaaatga ggttttetge 6420
gtocaaccag agaagggggg cogcaagcca gotogootta togtattoco agatttgggg 6480
gttcgtgtgt gcgagaaaat ggccctttac gatgtggtct ccaccctccc tcaggccgtg 6540
atgggctctt catacggatt ccaatactct cctggacage gggtcgagtt cctggtgaat 6600 gcctggaaag cgaagaaatg ccctatgggc ttcgcatatg acacccgctg ttttgactca 6660
acggtcactg agaatgacat ccgtgttgag gagtcaatct accaatgttg tgacttggcc 6720
cccgaagcca gacaggccat aaggtcgctc acagagcggc tttacatcgg gggccccctg 6780
actaattota aagggoagaa otgoggotat ogooggtgoo gogogagogg tgtactgacg 6840 accagotgog gtaatacoot cacatgttac ttgaaggoog otgoggootg togagotgog 6900
aagetecagg actgeacgat getegtatge ggagacgace ttgtegttat etgtgaaage 6960
geggggacce aagaggacga ggegageeta egggeettea eggaggetat gaetagatae 7020 tetgeecee etggggacce geccaaacca gaataegaet tggagttgat aacateatge 7080
tectecaatg tgteagtege geaegatgea tetggeaaaa gggtgtaeta teteaccegt 7140
gaccccacca ecceettge gegggetgeg tgggagacag etagacacae tecagteaat 7200
tectggetag geaacateat catgtatgeg eccaecttgt gggeaaggat gateetgatg 7260 acteattet tetecateet tetageteag gaacaacttg aaaaageeet agattgteag 7320
actcatttct tctccatcct tctagctcag gaacaacttg aaaaagccct agattgtcag
atctacgggg cctgttactc cattgagcca cttgacctac ctcagatcat tcaacgactc 7380
catggeetta gegeatttte acteeatagt tacteteeag gtgagateaa tagggtgget 7440
tcatgcctca ggaaacttgg ggtaccgccc ttgcgagtct ggagacatcg ggccagaagt 7500 gtccgcgcta ggctactgtc ccaggggggg agggctgcca cttgtggcaa gtacctcttc 7560
aactgggcag taaggaccaa geteaaacte acteeaatee eggetgegte ceagttggat 7620
```

		tggttacagc				
gcccgacccc	gctggttcat	gtggtgccta	ctcctacttt	ctgtaggggt	aggcatctat	7740
ctactcccca	accgatgaac	ggggagctaa	acactccagg	ccaataggcc	atcctgtttt	7800
		tttttttt				
tttttcctct	ttttttcctt	ttctttcctt	tggtggctcc	atcttagccc	tagtcacggc	7920
tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	7980
agatcaagt					•	7989

```
<210> 5
<211> 8649
<212> DNA
<213> Hepatitis C Virus
```

```
<400> 5
gecagecece gattggggge gacactecae catagateae teceetgtga ggaactaetg 60
tettcacgca gaaagegtet agecatggeg ttagtatgag tgtcgtgcag cetecaggae 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg tectitettg gateaacccg cteaatgeet ggagatttgg gegtgeecee 240
gegagactge tagecgagta gtgttgggte gegaaaggee ttgtggtaet geetgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcccat gattgaacaa gatggattgc 420
acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg gcacaacaga 480
caateggetg etetgatgee geogtgttee ggetgteage geaggggege eeggttettt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg getgetattg ggcgaagtge eggggeagga teteetgtea teteacettg 720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
cogaactgtt cgccaggctc aaggegegea tgcccgacgg cgaggatete gtegtgacce 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg 1140
ctcccqattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa eggttteect ctagegggat caatteegee ceteteeste ecceeceet 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tocaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgageatte ctaggggtet tteceetete gecaaaggaa tgeaaggtet gttgaatgte 1440
gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
togaggttaa aaaacgtota ggccccccga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatggaccg ggagatggca gcatcgtgcg gaggcgcggt tttcgtaggt 1860 ctgatactct tgaccttgtc accgcactat aagctgttcc tcgctaggct catatggtgg 1920 ttacaatatt ttatcaccag ggccgaggca cacttgcaag tgtggatccc ccccctcaac 1980
gttegggggg geegegatge egteateete etcaegtgeg egateeace agagetaate 2040 tttaceatea ceaaaatett getegeeata eteggteeae teatggtget eeaggetggt 2100
ataaccaaag tgccgtactt cgtgcgcgca cacgggctca ttcgtgcatg catgctggtg 2160
cggaaggttg ctgggggtca ttatgtccaa atggctctca tgaagttggc cgcactgaca 2220
ggtacgtacg tttatgacca tctcacccca ctgcgggact gggcccacgc gggcctacga 2280
gacettgegg tggcagttga gcccgtcgtc ttctctgata tggagaccaa ggttatcacc 2340
tggggggcag acaccgcggc gtgtggggac atcatcttgg gcctgcccgt ctccgcccgc 2400
agggggaggg agatacatct gggaccggca gacagccttg aagggcaggg gtggcgactc 2460
ctcgcgccta ttacggccta ctcccaacag acgcgaggcc tacttggctg catcatcact 2520
agecteacag geoggacag gaaccaggte gagggggagg tecaagtggt etceaeegea 2580 acacaatett teetggegae etgegteaat ggegtgtgtt ggaetgteta teatggtgee 2640
ggctcaaaga cccttgccgg cccaaagggc ccaatcaccc aaatgtacac caatgtggac 2700
caggaceteg teggetggea agegeeeece ggggegegtt cettgacace atgeacetge 2760
ggcagctcgg acctttactt ggtcacgagg catgccgatg tcattccggt gcgccggcgg 2820
ggcgacagca gggggagcct actctcccc aggcccgtct cctacttgaa gggctcttcg 2880 ggcggtccac tgctctgcc ctcggggcac gctgtgggca tctttcgggc tgccgtgtgc 2940
accegagggg tigegaagge ggtggacttt gtaceegteg agtetatgga aaccactatg 3000
cggtccccgg tcttcacgga caactcgtcc cctccggccg taccgcagac attccaggtg 3060
geceatetac aegecectae tggtagegge aagageacta aggtgeegge tgegtatgea 3120
gcccaagggt ataaggtgct tgtcctgaac ccgtccgtcg ccgccaccct aggtttcggg 3180 gcgtatatgt ctaaggcaca tggtatcgac cctaacatca gaaccggggt aaggaccatc 3240
accacgggtg cocccatcac gtactocacc tatggcaagt ttottgccga cggtggttgc 3300
tetggggggg cetatgacat cataatatgt gatgagtgec acteaactga ctegaccact 3360
atectgggca teggcaeagt eetggaeeaa geggagaegg etggagegeg aetegtegtg 3420
ctcgccaccg ctacgcctcc gggatcggtc accgtgccac atccaaacat cgaggaggtg 3480 gctctgtcca gcactggaga aatccccttt tatggcaaag ccatccccat cgagaccatc 3540
aaggggggga ggcacctcat tttctgccat tccaagaaga aatgtgatga gctcgccgcg 3600
```

```
aagetgteeg geeteggaet caatgetgta geatattace ggggeettga tgtateegte 3660
ataccaacta geggagaegt cattgtegta geaacggaeg ctetaatgae gggetttaee 3720
ggcgatttcg actcagtgat cgactgcaat acatgtgtca cccagacagt cgacttcage 3780
ctggacccga cetteaceat tgagacgacg accgtgecae aagacgeggt gteacgeteg 3840
cageggegag geaggactgg taggggeagg atgggeattt acaggtttgt gactecagga 3900
gaacggccct cgggcatgtt cgattcctcg gttctgtgcg agtgctatga cgcgggctgt 3960 gcttggtacg agctcacgcc cgccgagacc tcagttaggt tgcgggctta cctaaacaca 4020
ccagggttgc ccgtctgcca ggaccatctg gagttctggg agagegtctt tacaggcctc 4080
acceacatag acgeceattt ettgteecag actaageagg caggagacaa etteecetae 4140
ctggtagcat accaggetac ggtgtgcgcc agggetcagg ctccacctcc atcgtgggac 4200
caaatgtgga agtgtctcat acggctaaag cctacgctgc acgggccaac gcccctgctg 4260
tataggetgg gageegttea aaacgaggtt actaccacae acceeataae caaatacate 4320
atggcatgca tgtcggctga cctggaggtc gtcacgagca cctgggtgct ggtaggcgga 4380
gtectageag etetggeege gtattgeetg acaacaggea gegtggteat tgtgggeagg 4440 atcatettgt eeggaaagee ggeeateatt eeggacaggg aagteettta eegggagtte 4500
gatgagatgg aagagtgcgc ctcacacctc ccttacatcg aacagggaat gcagctcgcc 4560
gaacaattca aacagaaggc aatcgggttg ctgcaaacag ccaccaagca agcggaggct 4620
getgeteceg tggtggaate caagtggegg accetegaag cettetggge gaageatatg 4680
tggaatttca tcagcgggat acaatattta gcaggcttgt ccactctgcc tggcaacccc 4740
gegatageat cactgatgge atteacagee tetateacea gecegeteae cacceaacat 4800
accetectgt ttaacateet ggggggatgg gtggeegeee aacttgetee teccageget 4860 gettetgett tegtaggege eggeateget ggageggetg ttggeageat aggeettggg 4920
aaggtgettg tggataíttt ggeaggttat ggageagggg tggeaggege getegtggee 4980
tttaaggtca tgageggega gatgeeetee acegaggace tggttaaeet acteeetget 5040
atcetetece etggegeet agtegteggg gregtgtgeg cagegataet gegteggeae 5100 gtgggeecag gggagggge tgtgeagtgg atgaacegge tgatagegtt egettegegg 5160
ggtaaccacg totoccccac gcactatgtg cotgagageg acgetgeage acgtgtcact 5220
cagatectet etagtettae cateacteag etgetgaaga ggetteacea gtggateaac 5280
gaggactgct ccaegccatg ctccggctcg tggctaagag atgtttggga ttggatatgc 5340
acggtgttga ctgatttcaa gacctggctc cagtccaagc tcctgccgcg attgccggga 5400
gtccccttct tctcatgtca acgtgggtac aagggagtct ggcggggcga cggcatcatg 5460
caaaccacct gcccatgtgg agcacagate accggacatg tgaaaaacgg ttccatgagg 5520 atcgtggggc ctaggacctg tagtaacacg tggcatggaa cattccccat taacgcgtac 5580
accaegggee cetgeacgee etecceggeg ceaaattatt etagggeget gtggegggtg 5640
getgetgagg agtacgtgga ggttacgegg gtgggggatt tecaetacgt gaegggeatg 5700
accactgaca acgtaaagtg cccgtgtcag gttccggccc ccgaattctt cacagaagtg 5760
gatggggtge ggttgcacag gtacgctcca gcgtgcaaac ccctcctacg ggaggaggtc 5820
acattectgg tegggeteaa teaatacetg gttgggteae ageteecatg egageegaa 5880
coggacgtag cagigotoac ttocatgoto accgaeccet cocacattae ggeggagaeg 5940
gctaagcgta ggctggccag gggatetece eceteettgg ccageteate agetagecag 6000
ctgtctgcgc cttccttgaa ggcaacatgc actacccgtc atgactcccc ggacgctgac 6060
ctcatcgagg ccaacctcct gtggcggcag gagatgggcg ggaacatcac ccgcgtggag 6120 tcagaaaata aggtagtaat tttggactct ttcgagccgc tccaagcgga ggaggatgag 6180
agggaagtat ccgttccggc ggagatcctg cggaggtcca ggaaattccc tcgagcgatg 6240
cccatatggg cacgcccgga ttacaaccct ccactgttag agtcctggaa ggacccggac 6300
tacgtecete cagtggtaca egggtgteca ttgeegeetg ceaaggeee teegatacea 6360 eetecaegga ggaagaggae ggttgteetg teagaateta eegtgtette tgeettggeg 6420
gagetegeca caaagaeett eggeagetee gaategtegg eegtegacag eggeaeggea 6480
acggcctctc ctgaccagcc ctccgacgac ggcgacgcgg gatccgacgt tgagtcgtac 6540 tcctccatgc cccccttga gggggagccg ggggatcccg atctcagcga cgggtcttgg 6600 tctaccgtaa gcgaggagc tagtgaggac gtcgtctgct gctcgatgtc ctacacatgg 6660
acaggegeee tgateaegee atgegetgeg gaggaaacea agetgeeeat caatgeaetg 6720
agcaactett tgeteegtea ecacaacttg gtetatgeta caacateteg cagegeaage 6780
ctgcggcaga agaaggtcac ctttgacaga ctgcaggtcc tggacgacca ctaccgggac 6840
gtgctcaagg agatgaaggc gaaggcgtcc acagttaagg ctaaacttct atccgtggag 6900
gaagcetgta agetgaegee eccaeatteg geeagateta aatttggeta tggggeaaag 6960 gaegteegga acetateeag caaggeegtt aaceacatee geteegtgtg gaaggaettg 7020
ctggaagaca ctgagacacc aattgacacc accatcatgg caaaaaatga ggttttctgc 7080
gtccaaccag agaaggggg ccgcaagcca gctcgcctta tcgtattccc agatttgggg 7140
gttcgtgtgt gcgagaaaat ggccctttac gatgtggtct ccaccctccc tcaggccgtg
                                                                               7200
atgggctctt catacggatt ccaatactct cctggacage gggtcgagtt cctggtgaat 7260
geetggaaag egaagaaatg eectatggge ttegeatatg acaceegetg ttttgactea 7320
acggtcactg agaatgacat ccgtgttgag gagtcaatct accaatgttg tgacttggcc 7380
cccgaagcca gacaggccat aaggtcgctc acagagcggc tttacatcgg gggccccctg 7440
actaatteta aagggeagaa etgeggetat egeeggtgee gegegagegg tgtaetgaeg 7500 accagetgeg gtaataceet cacatgttae ttgaaggeeg etgeggeetg tegagetgeg 7560
aagetecagg actgeacgat getegtatge ggagacgace ttgtegttat etgtgaaage 7620
```

gcggggaccc	aagaggacga	ggcgagccta	cgggccttca	cggaggctat	gactagatac	7680
tetgececce	ctggggaccc	gcccaaacca	gaatacgact	tggagttgat	aacatcatgc	7740
tcctccaatg	tgtcagtcgc	gcacgatgca	tctggcaaaa	gggtgtacta	tctcacccgt	7800
gaccccacca	cccccttgc	gegggetgeg	tgggagacag	ctagacacac	tccagtcaat	7860
tcctggctag	gcaacatcat	catgtatgcg	cccaccttgt	gggcaaggat	gatcctgatg	7920
actcatttct	tctccatcct	tctagctcag	gaacaacttg	aaaaagccct	agattgtcag	7980
atctacgggg	cctgttactc	cattgagcca	cttgacctac	ctcagatcat	tcaacgactc	8040
catggcctta	gcgcattttc	actccatagt	tactctccag	gtgagatcaa	tagggtggct	8100
tcatgcctca	ggaaacttgg	ggtaccgccc	ttgcgagtct	ggagacatcg	ggccagaagt	8160
gtccgcgcta	ggctactgtc	ccaggggggg	agggctgcca	cttgtggcaa	gtacctcttc	8220
aactgggcag	taaggaccaa	gctcaaactc	actccaatcc	cggctgcgtc	ccagttggat	8280
					cctgtctcgt	8340
gcccgacccc	gctggttcat	gtggtgccta	ctcctacttt	ctgtaggggt	aggcatctat	8400
					atcctgtttt	8460
tttccctttt	ttttttttt	tttttttt	tttttttt	tttttttt	ttctcctttt	8520
tttttcctct	ttttttcctt	ttctttcctt	tggtggctcc	atcttagccc	tagtcacggc	8580
tagctgtgaa	aggtccgtga	gccgcttgac	tgcagagagt	gctgatactg	gcctctctgc	8640
agatcaagt						8649

```
<210> 6
<211> 8001
<212> DNA
<213> Hepatitis C Virus
<400> 6
gecagecece gattggggge gacaetecae catagateae teceetgtga ggaactaetg 60
tetteacgea gaaagegtet agecatggeg tragtatgag tgtegtgeag eetecaggae 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg teetteettg gateaacceg etcaatgeet ggagatttgg gegtgeecee 240
gcgagactgo tagoogagta gtgttgggto gogaaaggoo ttgtggtact gootgatagg 300 gtgcttgoga gtgccccggg aggtotogta gaccgtgoac catgagoacg aatootaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgeaggtte teeggeeget tgggtggaga ggetattegg etatgaetgg geacaacaga 480
caateggetg etetgatgee geegtgttee ggetgteage geaggggege eeggttettt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga teteetgtca teteacettg
                                                                              720
ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
cggctacctg cccattcgac caccaagega aacatcgcat cgagcgagca cgtactcgga 840
tggaagccgg tettgtegat caggatgate tggacgaaga gcatcagggg etegegecag 900
cogaactgtt egecaggete aaggegegea tgecegaegg egaggatete gtegtgaeee 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020 actgtggccg gctgggtgt gcggaccgct atcaggacat agcgttggct acccgtgata 1080 ttgctgaaga gcttggcgc gaatgggctg accgcttcct cgtgctttac ggtatcgccg 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320 tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcatte ctaggggtet tteceetete gecaaaggaa tgeaaggtet gttgaatgte 1440
gtgaaggaag cagtteetet ggaagettet tgaagacaaa caacgtetgt agegaeeett 1500
tgcaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgetttac atgtgtttag 1740
tegaggttaa aaaacgteta ggeececega accaegggga egtggtttte etttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcatcatca ctagcctcac aggccgggac aggaaccagg tcgaggggga ggtccaagtg 1920
gictccaccg caacacaatc titcciggcg acctgcgtca atggcgigig iiggactgic 1980 tatcatggtg ccggctcaaa gacccttgcc ggcccaaagg gcccaatcac ccaaatgtac 2040
accaatgigg accaggacet egicggeigg caagegeeee eeggggegeg tieetigaea 2100
ccatgcacct geggcagete ggacetttae ttggtcacga ggcatgcega tgtcattceg 2160 gtgcgcegge ggggcgacag cagggggage ctactetece ccaggecegt etcetacttg 2220 aagggetett egggeggtee actgetetge ecctegggge atgetgtggg catetttegg 2280
getgeegtgt geaccegagg ggttgegaag geggtggaet ttgtaccegt egagtetatg 2340
gaaaccacta tgeggteec ggtetteaeg gacaactegt eeeeteegge egtacegeag 2400 acatteeagg tggeeeatet acacgeeet actggtageg geaagageae taaggtgeeg 2460
getgegtatg cageccaagg gtataaggtg ettgteetga accegteegt egeegeeace 2520 etaggttteg gggegtatat gtetaaggea catggtateg accetaacat cagaaceggg 2580
gtagggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtggtt gctctggggg cgcctatgac atcataatat gtgatgagtg ccactcaact 2700
gactogacca ctatoctggg catoggoaca gtoctggaco aagoggagao ggotggagog 2760
egactegteg tgetegecae egetaegect eegggategg teaeegtgee acatecaaae 2820
ategaggagg tggetetgte cagcactgga gaaateeeet tttatggcaa agccateeee 2880
ategegacca teaagggggg gaggeaccte atttetgee attecaagaa gaaatgtgat 2940
gagetegeeg egaagetate eggeetegga eteaatgetg tageatatta eeggggeett 3000
gatgtatecg teataceaac tageggagae gteattgteg tageaacgga egetetaatg 3060
acgggettta eeggegattt egacteagtg ategactgea atacatgtgt cacceagaca 3120
gtegaettea geetggaece gaeetteace attgagaega egaeegtgee acaagaegeg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctcgggcatg ttcgattcct cggttctgtg cgagtgctat 3300
gacgeggget gtgcttggta egageteaeg ceegeegaga ceteagttag gttgeggget
                                                                              3360
tacctaaaca caccagggct geeegtetge caggaceate tggagttetg ggagagegte 3420
tttacaggee teacceacat agacgeceat ttettgteec agactaggea ggeaggagae 3480
aactteeeet aeetggtage afaceagget aeggtgtgeg eeagggetea ggeteeaeet 3540
```

ccatcgtggg accaaatgtg ggagtgtctc atacggctaa agcctacgct gcacgggcca 3600

```
acgecectge tgtatagget gggageegtt caaaacgagg ttactaceae acaceceata 3660
accaaataca tcatggcatg catgtcggct gacctggagg tcgtcacgag cacctgggtg 3720
ctggtaggcg gagtcctagc agctctggcc gcgtattgcc tgacaacagg cagcgtggtc 3780
attgtgggca ggatcatett gteeggaaag ceggecatea tteeegacag ggaagteett 3840
taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
atgcagctcg ccgaacaatt caaacagaag gcaatcgggt tgctgcaaac agccaccaag 3960
caageggagg etgetgetee egtggtggaa tecaagtgge ggaeeetega ageettetgg 4020 gegaageata tgtggaatt cateageggg atacaatatt tageaggett gtecaetetg 4080
cotggcaacc cogogatage atcactgatg gcattcacag cototatcac cagocogote 4140
accacccaac ataccctcct gtttaacatc ctggggggat gggtggccgc ccaacttgct 4200
cctcccagcg ctgcttctgc tttcgtaggc gccggcatcg ctggagcggc tgttggcagc 4260 ataggccttg ggaaggtgct tgtggatatt ttggcaggtt atggagcagg ggtggcaggc 4320
gegetegtgg cetttaaggt catgagegge gagatgeeet ceaeegagga eetggttaae 4380
ctactccctg ctatcctctc ccctggcgcc ctagtcgtcg gggtcgtgtg cgcagcgata 4440 ctgcgtcggc acgtgggccc aggggagggg gctgtgcagt ggatgaaccg gctgatagcg 4500 ttcgcttcgc ggggtaacca cgtctcccc acgcactatg tgtctgagag cgacgctgca 4560
geaegtgtea etcagateet etctagtett accareacte agetgetgaa gaggetteae 4620
cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680 gattggatat gcacggtgtt gactgatttc aagacctggc tccagtccaa gctcctgccg 4740
cgattgccgg gagtcccctt cttctcatgt caacgtgggt acaagggagt ctggcggggc 4800 gacggcatca tgcaaaccac ctgcccatgt ggagcacaga tcaccggaca tgtgaaaaac 4860
ggttccatga ggatcgtggg gcctaggacc tgtagtaaca cgtggcatgg aacattcccc 4920
attaacgcgt acaccacggg cccctgcacg ccctccccgg cgccaaatta ttctagggcg 4980 ctgtggcggg tggctgctga ggagtacgtg gaggttacgc gggtggggga tttccactac 5040
gtgacgggca tgaccactga caacgtaaag tgcccgtgtc aggttccggc ccccgaattc 5100
ttcacagaag tggatggggt gcggttgcac aggtacgctc cagcgtgcaa acccctccta 5160
cgggaggagg teacattect ggtegggete aateaatace tggttgggte acageteeca 5220
tgegggeeeg aaceggaegt ageagtgete acttecatge teacegaece eteceacatt 5280
acggeggaga eggetaageg taggetggee aggggatete ecceeteett ggeeagetea 5340
teagetagee agetgtetge geetteettg aaggeaacat geactaceeg teatgactee 5400
ccggacgctg acctcatcga ggccaacctc ctgtggcggc aggagatggg cgggaacatc 5460
accegegtgg agteagaaaa taaggtagta attttggaet etttegagee geteeaageg 5520
gaggaggatg agagggaagt atcogttoog goggagatoo tgoggaggto caggaaatto 5580
cctcgagcga tgcccatatg ggcacgcccg gattacaacc ctccactgtt agagtcctgg 5640
aaggacccgg actacgtccc tccagtggta cacgggtgtc cattgccgcc tgccaaggcc 5700
cctccgatac cacctccacg gagggagagg acggttgtcc tgtcagaatc taccgtgtct 5760
totgoottgg oggagotogo cacaaagaco ttoggoagot cogaatogto ggoogtogac 5820
ageggeaegg caaeggeete teetgaceag eesteegaeg aeggegaege gggateegae 5880
gttgagtcgt actoctccat geoccccctt gagggggage egggggatee egateteage 5940
gacgggtctt ggtctaccgt aagcgaggag gctagtgagg acgtcgtctg ctgctcgatg 6000
tectacacat ggacaggege cetgateacg ceatgegetg eggaggaaac caagetgeee 6060
gtcaatgcac tgagcaactc tttgctccgt caccacaact tggtctatgc tacaacatct 6120
cgcagcgcaa gcctgcggca gaagaaggtc acctttgaca gactgcaggt cctggacgac 6180
cactaceggg acgtgctcaa ggagatgaag gegaaggegt ccacagttaa ggetaaactt 6240
ctatecgtgg aggaageetg taagetgaeg cececacatt eggeeagate taaatttgge 6300
tatggggcaa aggacgtccg gaacctatcc agcaaggccg ttaaccacat ccgctccgtg 6360
tggaaggact tgctggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
gaggttttet gegtecaace agagaagggg ggeegeaage eagetegeet tategtatte 6480
ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggt ctccaccctc 6540
cctcaggccg tgatgggctc ttcatacgga ttccaatact ctcctggaca gcgggtcgag 6600 ttcctggtga atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacacccgc 6660 tgttttgact caacggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt 6720
tgtgacttgg cccccgaage cagacaggee ataaggtege tcacagageg getttacate 6780
gggggccccc tgactaattc taaagggcag aactgcggct atcgccggtg ccgcgcgagc 6840
ggtgtactga cgaccagetg eggtaatace etcacatgtt atttgaagge egetgeggee 6900
tgtcgagctg cgaagctcca ggactgcacg atgctcgtat gcggagacga ccttgtcgtt 6960
atetgtgaaa gegeggggae ecaagaggae gaggegagee taegggeett caeggagget 7020
atgactagat actotgocco coctggggac cogoccaaac cagaatacga cttggagttg 7080 ataacatcat gotoctccaa tgtgtcagtc gogcacgatg catctggcaa aagggtgtac 7140 tatctcaccc gtgaccccac caccccctt gogcgggctg cgtgggagac agctagacac 7200
actocagtoa attoctggot aggoaacato atcatgtatg ogcocacott gtgggcaagg 7260
atgatectga tgaeteattt ettetecate ettetagete aggaacaact tgaaaaagee 7320
ctagattgtc agatctacgg ggcctgttac tccattgagc cacttgacct acctcagatc 7380 attcaacgac tccatggcct tagcgcattt tcactccata gttactctcc aggtgagatc 7440
aatagggtgg cttcatgcct.caggaaactt ggggtaccgc ccttgcgagt ctggagacat 7500
cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560 aagtacetet tcaactgggc agtaaggacc aagetcaaac tcactccaat eccggetgeg 7620
```

tcccagttgg	atttatccag	ctggttcgtt	gctggttaca	gcgggggaga	catatatcac	7680
agcctgtctc	qtqcccqacc	ccgctggttc	atgtggtgcc	tactcctact	ttctgtaggg	7740
gtaggcatct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctgtt	tttttccctt	ttttttttc	tttttttt	tttttttt	tttttttt	7860
ttttctcctt						
cctagtcacg						
	gcagatcaag			•		8001

```
<211> 11076
<212> DNA
<213> Hepatitis C Virus
<400> 7
gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
tetteacgea gaaagegtet agecatggeg ttagtatgag tgtegtgeag cetecaggae 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg teetttettg gateaacccg etcaatgeet ggagatttgg gegtgeece 240 gegagactge tageegagta gtgttgggte gegaaaggee ttgtggtaet geetgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgcaggttc teeggeeget tgggtggaga ggetattegg ctatgaetgg gcacaacaga 480
caateggetg etetgatgee geegtgttee ggetgteage geaggggege eeggttettt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggacig getgetatig ggegaagige eggggeagga teteetgiea teteacetig 720 eteetgeega gaaagtatee ateatggetg atgeaatgeg geggetgeat acgettgate 780
eggetacetg cecattegae caccaagega aacategeat egagegagea egtactegga 840
tggaageegg tettgtegat eaggatgate tggaegaaga geateagggg etegegeeag 900
cegaactgtt egecaggete aaggegegea tgeeegaegg egaggatete gtegtgaeee 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg 1140 ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200 cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tocaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggtct ttcccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440 gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500
tgcaggcage ggaaccecce acctggcgae aggtgcetet geggeeaaaa gecaegtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680 gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
tcgaggttaa aaaacgtcta ggccccccga accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcggtgg agtttacctg 1920 ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccg ggtaccctg gcccctctat ggcaatgagg gcttggggtg ggcaggatgg 2100 ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160 cgcaatttgg gtaaggtcat cgatacctc acgtgcggct tcgccgatct catggggtac 2220
attecgeteg teggegeee eetaggggge getgeeaggg ceetggegea tggegteegg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340 ttccttttgg ctttgctgt ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400 gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
geggacatga teatgeatae eecegggtge gtgeeetgeg ttegggagaa caacteetee 2520
cgctgctggg tagcgctcac tcccacgctc gcggccagga acgctagcgt ccccactacg 2580
acgatacgac gecatgtega tttgetegtt ggggeggetg etetetgete egetatgtac 2640 gtgggagate tetgeggate tgttteete gtegeceage tgtteacett etegeetege 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgcag 2820
ttactccgga tcccacaagc tgtcgtggat atggtggcgg gggcccattg gggagtccta 2880 gcgggccttg cctactattc catggtgggg aactgggcta aggttctgat tgtgatgcta 2940
ctctttgccg gcgttgacgg gggaacctat gtgacagggg ggacgatggc caaaaacacc 3000
ctcgggatta cgtccctctt ttcacccggg tcatcccaga aaatccagct tgtaaacacc 3060
aacggcaget ggcacatcaa caggactgcc etgaactgca atgactccct caacactggg 3120
ttoottgotg ogotgttota ogtgoacaag ttoaactoat otggatgooc agagogoatg 3180
gccagctgca gccccatcga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240
tcacacaget eggaceagag geettattgt tggcactaeg caceeeggee gtgeggtate 3300
gtacccgcgg cgcaggtgtg tggtccagtg tactgcttca ccccaagecc tgtcgtggtg 3360 gggacgaccg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gccgccgcaa ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt teaccaagae gtgegggge eeceegtgta acateggggg gateggeaat 3540 aaaaeettga eetgeeceae ggaetgette eggaageaee eegaggeeae ttacaccaag 3600
```

```
tgtggttcgg ggccttggtt gacacccaga tgcttggtcc actacccata caggctttgg 3660 cactacccct gcactgtcaa ctttaccatc ttcaaggtta ggatgtacgt ggggggagtg 3720
gagcacaggo togaagcogo atgcaattgg actogaggag agogttgtaa cotggaggac 3780
agggacagat cagagettag ecegetgetg etgtetacaa eggagtggea ggtattgeee
tgttccttca ccaccctacc ggctctgtcc actggtttga tccatctcca tcagaacgtc
                                                                                  3900
gtggacgtac aatacctgta cggtataggg tcggcggttg tctcctttgc aatcaaatgg 3960
gagtatgtcc tgttgctctt ccttcttctg gcggacgcgc gcgtctgtgc ctgcttgtgg 4020
atgatgctgc tgatagctca agctgaggcc gccctagaga acctggtggt cctcaacgcg 4080
gcatecgtgg ccggggcgca tggcattete teetteeteg tgttettetg tgetgeetgg 4140 tacateaagg gcaggetggt ecetggggcg gcatatgeee tetaeggcgt atggcegeta 4200 eteetgetee tgetggegtt accaccacga gcatacgcca tggaccggga gatggcagca 4260
tegtgeggag gegeggtttt egtaggtetg atactettga cettgteace geactataag 4320
ctgtteeteg ctaggeteat atggtggtta caatattta teaccagge cgaggeacae 4380
ttgcaagtgt ggatccccc cctcaacgtt cgggggggcc gcgatgccgt catcctcctc 4440
acgtgcgcga tccacccaga gctaatcttt accatcacca aaatcttgct cgccatactc 4500
ggtccactca tggtgctcca ggctggtata accaaagtgc cgtacttcgt gcgcgcacac 4560
gggctcattc gtgcatgcat gctggtgcgg aaggttgctg ggggtcatta tgtccaaatg 4620 gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct caccccactg 4680
cgggactggg cccacgeggg cctacgagac cttgeggtgg cagttgagcc cgtcgtcttc 4740
totgatatgg agaccaaggt tatcacctgg ggggcagaca ccgcggcgtg tggggacatc 4800
atcitgggcc tgcccgtctc cgcccgcagg gggagggaga tacatctggg accggcagac 4860 agccttgaag ggcagggtg gcgactcctc gcgcctatta cggcctactc ccaacagacg 4920
cgaggeetae ttggetgeat cateactage ctcacaggee gggacaggaa ccaggtegag 4980
ggggaggtcc aagtggtctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
gtgtgttgga ctgtctatca tggtgccggc tcaaagaccc ttgccggccc aaagggccca 5100
atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gcccccggg 5160
gegegtteet tgacaccatg cacetgegge ageteggace tttacttggt cacgaggeat 5220
geogatgtea tteeggtgeg eeggeggge gacageaggg ggageetact eteeceagg 5280
cccgtctcct acttgaaggg ctcttcgggc ggtccactgc tctgcccctc ggggcatgct 5340 gtgggcatct ttcgggctgc cgtgtgcacc cgaggggttg cgaaggcggt ggactttgta 5400
cocglogagt ctatggaaac cactatgegg tecceggtet teaeggacaa etegteeeet 5460
coggoogtac ogcagacatt coaggtggcc catchacacg cocctactgg tagoggcaag 5520
ageactaagg tgeeggetge gtatgeagee caagggtata aggtgettgt cetgaaceeg 5580 teegtegeeg ceaccetagg ttteggggeg tatatgteta aggeacatgg tategaeeet 5640
aacatcagaa ccggggtagg gaccatcacc acgggtgccc ccatcacgta ctccacctat 5700
ggcaagtttc ttgccgacgg tggttgctct gggggcgcct atgacatcat aatatgtgat 5760
gagtgccact caactgactc gaccactate ctgggcatcg gcacagtcct ggaccaagcg 5820 gagacggctg gagcgcgact cgtcgtgctc gccaccgcta cgcctccggg atcggtcacc 5880
gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
ggcaaagcca tccccatcgc gaccatcaag ggggggaggc acctcatttt ctgccattcc 6000
aagaagaaat gtgatgaget egeegegaag etateeggee teggaeteaa tgetgtagea 6060 tattaeeggg geettgatgt ateegteata ceaactageg gagaegteat tgtegtagea 6120 aeggaegete taatgaeggg etttaeegge gatttegaet eagtgatega etgeaataca 6180
tgtgtcaccc agacagtcga cttcagcctg gacccgacct tcaccattga gacgacgacc 6240
gtgccacaag acgcggtgtc acgctcgcag cggcgaggca ggactggtag gggcaggatg 6300
ggcatttaca ggtttgtgac tccaggagaa cggccctcgg gcatgttcga ttcctcggtt 6360
ctgtgcgagt gctatgacgc gggctgtgct tggtacgagc tcacgcccgc cgagacctca 6420
gttaggttgc gggcttacct aaacacacca gggctgcccg tctgccagga ccatctggag 6480
ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagact 6540
aggcaggcag gagacaactt cccctacetg gtagcatacc aggetacggt gtgcgccagg 6600
gctcaggctc cacctccatc gtgggaccaa atgtgggagt gtctcatacg gctaaagcct 6660
acgctgcacg ggccaacgcc cctgctgtat aggctgggag ccgttcaaaa cgaggttact 6720
accacacac ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780 acgagcacct gggtgctggt aggcggagtc ctagcagctc tggccgcgta ttgcctgaca 6840
acaggcageg tggtcattgt gggcaggate atettgteeg gaaageegge cateatteee 6900
gacagggaag teetttaeeg ggagttegat gagatggaag agtgegeete acaceteeet 6960
tacategaac agggaatgca getegeegaa caatteaaac agaaggeaat egggttgetg 7020
caaacagcca ccaagcaagc ggaggctgct gctcccgtgg tggaatccaa gtggcggacc 7080
ctcgaagcct tctgggcgaa gcatatgtgg aatttcatca gcgggataca atatttagca 7140
ggettgteca etetgeetgg caaccegeg atagcateae tgatggcatt cacagcetet 7200
atcaccagee egeteaceae ecaacatace etectgttta acatectggg gggatgggtg 7260
geogeocaae tigeteetee cagegotget tetgetiteg taggegoogg categotgga 7320
geggetgttg geageatagg eettgggaag gtgettgtgg atattttgge aggttatgga 7380
gcaggggtgg caggcgcgct cgtggccttt aaggtcatga gcggcgagat gccctccacc 7440
gaggacetgg ttaacetact ecctgetate eteteceetg gegeeetagt egteggggte 7500 gtgtgegeag egatactgeg teggeacgtg ggeeeagggg agggggetgt geagtggatg 7560 aaceggetga tagegttege ttegegggt aaceaegtet ecceeaegea etatgtgtet 7620
```

```
gagagegaeg etgeageaeg tgteaeteag atecteteta gtettaceat cacteagetg 7680
ctgaagaggc ttcaccagtg gatcaacgag gactgctcca cgccatgctc cggctcgtgg 7740
ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggctccag 7800
tocaagetee tgeegegatt geegggagte ecettettet eatgteaacg tgggtacaag 7860 ggagtetgge ggggegaegg catcatgeaa accasetgee eatgtggage acagateace 7920
ggacatgtga aaaacggttc catgaggatc gtggggccta ggacctgtag taacacgtgg 7980
catggaacat tececattaa egegtacace aegggeeest geacgeeete eeeggegesa 8040
aattatteta gggegetgtg gegggtgget getgaggagt aegtggaggt taegegggtg 8100 ggggatttee actaegtgae gggeatgaee aetgaeaaeg taaagtgeee gtgteaggtt 8160
ecggeeeecq aattetteae agaagtggat ggggtgeggt tgeacaggta egeteeageg 8220
tgcaaaccc tcctacgga ggaggtcaca ttcctggtcg ggctcaatca atacctggtt 8280 gggtcacagc tcccatgcgg gcccgaaccg gacgtagcag tgctcacttc catgctcacc 8340 gacccctcc acattacggc ggagacggct aagcgtaggc tggccagggg atctccccc 8400
teettegeca geteateage tagecagetg tetgegeett cettgaagge aacatgeact 8460
acceptcate actecoegga egetgaeete ategaggeea aceteetgtg geggeaggag 8520
atgggcggga acatcacccg cgtggagtca gaaaataagg tagtaatttt ggactctttc 8580 gagccgctcc aagcggagga ggatgagagg gaagtatccg ttccggcgga gatcctgcgg 8640
aggtecagga aatteceteg agegatgeee atatgggeae geeeggatta caacceteea 8700
ctgttagagt cctggaagga cccggactac gtccctccag tggtacacgg gtgtccattg 8760 ccgcctgcca aggccctcc gataccacct ccacggaggg agaggacggt tgtcctgtca 8820 gaatctaccg tgtcttctgc cttggcggag ctcgccacaa agaccttcgg cagctccgaa 8880
tegteggeeg tegacagegg caeggeaacg geeteteetg accageeete egacgaegge 8940
gacgegggat cegacgttga gtegtactee tecatgeece ceettgaggg ggageegggg 9000
gatocogato toagogacog gtottogtot acogtaagog aggaggotag tgaggacoto 9060
gtotgotgot ogatgtocta cacatggaca ggogocotga toacgccatg cgctgoggag 9120
gaaaccaage tgecegteaa tgeactgage aactetttge teegteacca caacttggte 9180
tatgctacaa catctcgcag cgcaagcctg cggcagaaga aggtcacctt tgacagactg 9240
caggteetgg acgaccacta eegggaegtg etcaaggaga tgaaggegaa ggegteeaca 9300
gttaaggcta aacttetate egtggaggaa geetgtaage tgaegeeece acatteggee 9360
agatetaaat ttggetatgg ggeaaaggae gteeggaace tateeageaa ggeegttaae 9420
cacatcogct cogtgtggaa ggacttgctg gaagacactg agacaccaat tgacaccacc 9480 atcatggcaa aaaatgaggt tttctgcgtc caaccagaga aggggggccg caagccagct 9540
cgccttatcg tattcccaga tttgggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600
gtggtctcca ccctccctca ggccgtgatg ggctcttcat acggattcca atactctcct 9660 ggacagcggg tcgagttcct ggtgaatgcc tggaaagcga agaaatgccc tatgggcttc 9720
gcatatgaca cocgetgttt tgactcaacg gtcactgaga atgacatecg tgttgaggag 9780
tcaatctacc aatgttgtga cttggccccc gaagccagac aggccataag gtcgctcaca 9840
gageggettt acateggggg ecceetgaet aattetaaag ggeagaaetg eggetatege 9900
cggtgccgcg cgagcggtgt actgacgacc agctgcggta ataccctcac atgttatttg 9960
aaggeegetg eggeetgteg agetgegaag etecaggaet geaegatget egtatgegga 10020
gacgacettg tegttatetg tgaaagegeg gggacecaag aggacgagge gageetaegg 10080
gccttcacgg aggctatgac tagatactct gcccccctg gggacccgcc caaaccagaa 10140
tacgacttgg agttgataac atcatgctcc tccaatgtgt cagtcgcgca cgatgcatct 10200
ggcaaaaggg tgtactatct cacccgtgac cccaccaccc cccttgcgcg ggctgcgtgg 10260
gagacageta gacacactee agteaattee tggetaggea acateateat gtatgegeee 10320
accttgtggg caaggatgat cetgatgact catttettet ceatecttet agetcaggaa 10380
caacttgaaa aageeetaga ttgtcagate taeggggeet gttaeteeat tgageeactt 10440 gaeetaeete agateattea aegaeteeat ggeettageg catttteaet ceatagttae 10500
totocaggtg agatcaatag ggtggottca tgcctcagga aacttggggt accgccttg 10560
cgagtetgga gacateggge cagaagtgte egegetagge taetgteeea gggggggagg 10620
gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680 ccaatcccgg ctgcgtcca gttggattta tccagctggt tcgttgctgg ttacagcggg 10740 ggagacatat atcacagcct gtctcgtgcc cgaccccgct ggttcatgtg gtgcctactc 10800
tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
agagagtgct gatactggcc tctctgcaga tcaagt
```

```
<211> 8001
<212> DNA
<213> Hepatitis C Virus
<400> 8
gecagecece gattggggge gacactecae catagateae teceetgtga ggaactaetg 60
tettéacqca gaaagegtet agecatggeg ttagtatgag tgtegtgeag cetecaggae 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaceggg teetttette gateaacceg etcaatgeet ggagatttgg gegtgeecee 240 gegagaetge tageegagta gtgttgggte gegaaaggee ttgtggtaet geetgatagg 300
gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgeaggtte teeggeeget tgggtggaga ggetattegg etatgaetgg geacaacaga 480
caateggetg etetgatgee geegtgttee ggetgteage geaggggege ceggttettt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttgtc actgaagcgg 660
gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720 ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780 cggctacctg cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga 840
tggaagccg tcttgtcgat caggatgatc tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactgtt cgccaggete aaggegegea tgcccgaegg cgaggatete gtcgtgaeee 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct accegtgata 1080
ttgctgaaga gettggegge gaatgggetg accepttect egtgetttac ggtategeeg 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200
cagaccacaa eggttteeet etagegggat caatteegee ceteteeete eccececet 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tocaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380
acgagcattc ctaggggtct ttcccctctc gccaaaggaa tgcaaggtct gttgaatgtc 1440
gtgaaggaag cagtteetet ggaagettet tgaagacaaa caacgtetgt agegaeeett 1500 tgeaggeage ggaaceeec acetggegae aggtgeetet geggeeaaaa geeaegtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740
togaggttaa aaaacgtota ggooccocga accacgggga cgtggtttto ctttgaaaaa 1800
cacgataata ccatggegec tattacggec tacteceaac agaegegagg cetaettgge 1860
tgcatcatca ctagcctcac aggccgggac aggaaccagg tcgaggggga ggtccaagtg 1920
gtotocaccg caacacaatc tttoctggcg acctgcgtca atggcgtgtg ttggactgtc 1980 tatcatggtg ccggctcaaa gacccttgcc ggcccaaagg gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacct cgtcggctgg caagcgcccc ccggggcgcg ttccttgaca 2100
ccatgcacct geggeagete ggacetttae ttggteaega ggeatgeega tgteatteeg 2160
gtgcgccggc ggggcgacag cagggggagc ctactctccc ccaggcccgt ctcctacttg 2220
aagggetett egggeggtee aetgetetge eeetegggge aegetgtggg catetttegg 2280
gctgccgtgt gcacccgagg ggttgcgaag gcggtggact ttgtacccgt cgagtctatg 2340
gaaaccacta tgcggtcccc ggtcttcacg gacaactcgt cccctccggc cgtaccgcag 2400 acattccagg tggcccarct acacgccct actggtagcg gcaagagcac taaggtgccg 2460
gctgcgtatg cagcccaagg gtataaggtg cttgtcctga acccgtccgt cgccgccacc 2520
ctaggtttcg gggcgtatat gtctaaggca catggtatcg accctaacat cagaatcggg 2580 gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640 gacggtggtt gctctgggg cgcctatgac atcataatat gtgatgagtg ccactcaact 2700 gactcgacca ctatcctggg catcggcaca gtcctggacc aagcggagga ggcttgaagcg 2760
cgactcgtcg tgctcgccac cgctacgcct ccgggatcgg tcaccgtgcc acatccaaac 2820
atcgaggagg tggctctgtc cagcactgga gaaatcccct tttatggcaa agccatcccc 2880
atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
gagetegeeg egaagetgte eggeetegga eteaatgetg tageatatta eeggggeett 3000
gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaatg 3060
acgggcttta ccggtgactt cgactcagtg atcgactgca atacatgtgt cacccagaca 3120
gtcgacttca gcctggaccc gaccttcacc attgagacga cgaccgtgcc acaagacgcg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctcgggcatg ttcgattcct cggttctgtg cgagtgctat 3300
gacgeggget gtgcttggta egageteaeg eeegeegaga eeteagttag gttgeggget 3360
tacctaaaca caccagggtt geeegtetge caggaceate tggagttetg ggagagegte 3420 tttacaggee teacceacat agacgeecat ttettgteec agactaagea ggeaggagae 3480
aacttcccct acctggtage ataccagget acggtgtgcg ccagggetca ggctccacct 3540
ccatcgtggg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacgggcca 3600
```

```
acgecectge tgtatagget gggageegtt caaaacgagg ttactaceae acaccecata 3660
accaaataca tcatggcatg catgtcggct gacctggagg tcgtcacgag cacctgggtg 3720
ctggtaggeg gagtectage agetetggee gegtattgee tgacaacagg cagegtggte 3780 attgtgggea ggateatett gteeggaaag ceggeeatea tteeegacag ggaagteett 3840
tacoggagt togatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
atgcagctcg ccgaacaatt caaacagaag gcaatcgggt tgctgcaaac agccaccaag 3960
caageggagg etgetgetee egtggtggaa tecaagtgge ggaceatega ageettetgg 4020 gegaageata tgtggaattt caleageggg atacaatatt tageaggett gtecaetetg 4080
cotggcaaco cogogatago atcactgatg goattcacag cototatoac cagooogoto 4140
accacccaac ataccctcct gtttaacatc ctggggggat gggtggccgc ccaacttgct 4200
cctcccagcg ctgcttctgc tttcgtaggc gccggcatcg ctggagcggc tgttggcagc 4260 ataggccttg ggaaggtgct tgtggatatt ttggcaggtt atggagcagg ggtggcaggc 4320
gegetegtgg eetttaaggt catgagegge gagatgeest ceaeegagga eetggttaac 4380
ctactccctg ctatcctctc ccctggcgcc ctagtcgtcg gggtcgtgtg cgcagcgata 4440
ctgcgtcggc acgtgggccc aggggagggg gctgtgcagt ggatgaaccg gctgatagcg 4500 ttcgcttcgc ggggtaacca cgtctcccc acgcactatg tgcctgagag cgacgctgca 4560
geaegtgtea eteagateet etetagtett accateacte agetgetgaa gaggetteae 4620
cagtggatca acgaggactg ctccacgcca tgctccggct cgtggctaag agatgtttgg 4680
gattggatat gcacggtgtt gactgattte aagacetgge tecagtecaa geteetgeeg 4740 egattgeegg gagteeeett etteteatgt caacgtgggt acaagggagt etggegggge 4800
gacggcatca tgcaaaccac ctgcccatgt ggggcacaga tcaccggaca tgtgaaaaac 4860
ggttccatga ggatcgtggg gcctaggacc tgtagtaaca cgtggcatgg aacattcccc 4920
attaacgcgt acaccacggg cocctgcacg coctccccgg cgccaaatta ttctagggcg 4980
ctgtggcggg tggctgctga ggagtacgtg gaggttacgc gggtgggggga tttccactac 5040
gtgacgggca tgaccactga cgacgtaaag tgcccgtgte aggttccggc ccccgaattc 5100
ttcacagaag tggatggggt gcggttgcac aggtacgctc cagcgtgcaa acccctccta 5160
cgggaggagg tcacattcct ggtcgggctc aatcaatacc tggttgggtc acagctccca 5220
tgcgagcccg aaceggatgt agcagtgcte acttecatge tcacegacce eteccacatt 5280
acggcggaga cggctaagcg taggctggcc aggggatotc ctccccctt ggccagctca 5340
tragetages agetgtetge geetteettg aaggeaacat geactacecg teatgactee 5400
ccggacgctg acctcatcga ggccaacctc ctgtggcggc aggagatggg cgggaacatc 5460
accegegtgg agteagaaaa taaggtagta attttggaet etttegagee geteeaageg 5520
gaggaggatg agagggaagt atccgttccg gcggagatcc tgcggaggtc caggaaattc 5580
cetegagega tgeceatatg ggeacgeeeg gattacaace etceaetgtt agagteetgg 5640 aaggaeeegg actacgteee tecagtggta caegggtgte cattgeegee tgecaaggee 5700
cotocgatac cacetteacg gaggaagagg acggttgtcc tgtcagaatc taccgtgtct 5760
totgeottag eggagetege cacagagace tteggeaget eegaategte ggeegtegae 5820
ageggeaegg caaeggeete teetgaeeag eeeteegaeg aeggegaege gggateegae 5880
gttgagtcgt actectecat gececectt gagggggage egggggatee egateteage 5940 gaegggtett ggtetaeegt aagegaggag getagtgagg aegtegtetg etgetegatg 6000
tectacaeat ggaeaggege cetgateaeg ceatgegetg eggaggaaae caagetgeee 6060
atcaatgcac tgagcaactc tttgctccgt caccacaact tggtctatgc tacaacatct 6120
cgcagcgcaa gcctgcggca gaagaaggtc acctttgaca gactgcaggt cctggacgac 6180
cactaceggg aegtgeteaa ggagatgaag gegaaggegt ceacagttaa ggetaaactt 6240
ctatccgtgg aggaagectg taagetgacg ecceacatt eggecagate taaatttgge 6300
tatggggcaa aggacgtccg gaacetatec agcaaggccg ttaaccacat ccgctccgtg 6360 tggaaggact tgctggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420 gaggttttct gcgtccaacc agagaagggg ggccgcaagc cagctcgcct tatcgtattc 6480
ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggt ctccaccctc 6540
ceteaggeeg tgatgggete tteataegga ttecaataet eteetggaea gegggtegag 6600
ttcctggtga atgcctggaa agcgaagaaa tgccctatgg gcttcgcata tgacacccgc 6660 tgttttgact caacggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt 6720
tgtgacttgg cccccgaagc cagacaggcc ataaggtcgc tcacagagcg gctttacatc 6780
gggggccccc tgactaattc taaagggcag aactgcggct atcgccggtg ccgcgcgagc 6840
ggtgtactga cgaccagctg cggtaatacc ctcacatgtt acttgaaggc cgctgcggcc 6900
tgtcgagctg cgaagctcca ggactgcacg atgctcgtat gcggagacga ccttgtcgtt 6960
atctgtgaaa gcgcggggac ccaagaggac gaggcgagcc tacgggcctt cacggaggct 7020
atgactagat actetgeece ecetggggae ecgeceaaac eagaataega ettggagttg 7080 ataacateat geteeteeaa tgtgteagte gegeaegatg eatetggeaa aagggtgtae 7140 tateteace gtgaceceae eaceeeett gegegggetg egtgggagae agetagaeae 7200
actocagtoa attoctggot aggoaacato atcatgtatg ogcocacott gtgggcaagg 7260
atgatectga tgaeteattt etteteeate ettetagete aggaacaact tgaaaaagee 7320
ctagattgtc agatctacgg ggcctgttac tccattgagc cacttgacct acctcagatc 7380 attcaacgac tccatggcct tagcgcattt tcactccata gttactctcc aggtgagatc 7440
aatagggtgg cttcatgcct caggaaactt ggggtaccgc ccttgcgagt ctggagacat 7500
cgggccagaa gtgtccgcgc taggctactg tcccaggggg ggagggctgc cacttgtggc 7560 aagtacctct tcaactgggc agtaaggacc aagctcaaac tcactccaat cccggctgcg 7620
```

tcccagttgg atttatcc	ag ctggttcgtt	gctggttaca	gcgggggaga	catatatcac	7680
agcctgtctc gtgcccga	cc ccgctggttc	atgtggtgcc	tactcctact	ttctgtaggg	7740
gtaggcatct atctacto	cc caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctgtt tttttccc	tt ttttttttc	tttttttt	ttttttttt	ttttttttt	7860
ttttctcctt ttttttc	ct cttttttcc	ttttctttcc	tttggtggct	ccatcttagc	7920
cctagtcacg gctagctg	tg aaaggtccgt	gagecgettg	actgcagaga	gtgctgatac	7980
tggcctctct gcagatca	ag t	, , , , , , , , , , , , , , , , , , , ,			8001

```
<211> 11076
 <212> DNA
 <213> Hepatitis C Virus
 <400> 9
 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60
 tetteaegea gaaagegtet ageeatggeg ttagtatgaq tgtegtgeag eeteeaggae 120
coccctcc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180 gacgaccggg tootttottg gatcaacccg ctcaatgcot ggagatttgg gcgtgccccc 240 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtggtact gcctgatagg 300
 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360
 ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
 acgeaggtte teeggeeget tgggtggaga ggetattegg etatgaetgg geacaacaga 480
caateggetg ctetgatgee geogtgttee ggetgteage geaggggege ceggttettt 540
 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
cgtggctggc cacgacggc gttccttgcg cagctgtgt cgacgttgtc actgaagcgg 660 gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca tctcaccttg 720 ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780
eggetacetg cecattegae caceaagega aacategeat egagegagea egtactegga 840°
tggaagccgg tottgtcgat caggatgate tggacgaaga gcatcagggg etegegecag 900 ccgaactgtt egecaggete aaggegegea tgeecgacgg egaggatete gtegtgaeee 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gettggege gaatgggetg acceptitect egtgetttae ggtategeeg 1140 etceegatte geagegeate geettetate geettettga egagttette tgagtttaaa 1200
cagaccacaa eggitteeet etagegggat caarteegee eeteteeete eececeeet 1260
aacgttactg gccgaagecg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320 tecaccatat tgccgtettt tggcaatgtg agggcccgga aacetggccc tgtettettg 1380 acgagcatte ctaggggtet ttecectete gccaaaggaa tgcaaggtet gttgaatgte 1440
gtgaaggaag cagtteetet ggaagettet tgaagacaaa caacgtetgt agegaccett 1500
tgcaggcagc ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgetttae atgtgtttag 1740
tegaggttaa aaaaegteta ggeeeeega aeeaegggga egtggtttte etttgaaaaa 1800
cacgataata ccatgggcac gaateetaaa eeteaaagaa aaaccaaacg taacaccaac 1860
cgccgcccac aggacgtcaa gttcccgggc ggtggtcaga tcgtcggtgg agtttacctg 1920
tigeogegea ggggeeceag gttgggtgtg egegegaeta ggaagaette egageggteg 1980
caacctcgtg gaaggcgaca acctatcccc aaggctcgcc agcccgaggg tagggcctgg 2040
gctcagcccg ggtacccctg gcccctctat ggcaatgagg gcttggggtg ggcaggatgg 2100 ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160
cgcaatttgg gtaaggtcat cgataccetc acgtgcggct tcgccgatct catggggtac 2220
attecgeteg teggegeee cetaggggge getgecaggg ceetggegea tggegteegg 2280
gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340 ttccttttgg ctttgctgtc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400
gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460
geggaeatga teatgeatae eecegggtge gtgeeetgeg ttegggagaa caacteetee 2520
cgctgctggg tagcgctcac tcccacgctc gcggccagga acgctagcgt ccccactacg 2580 acgatacgac gccatgtcga tttgctcgtt ggggcggctg ctctctgctc cgctatgtac 2640
gtgggagate tetgeggate tgtttteete gtegeecage tgttcaeett etegeetege 2700
cggcacgaga cagtacagga ctgcaattgc tcaatatatc ccggccacgt gacaggtcac 2760
cgtatggctt gggatatgat gatgaactgg tcacctacag cagccctagt ggtatcgcag 2820
ttactccgga tcccacaagc tgtcgtggat atggtggcgg gggcccattg gggagtccta 2880
gegggeettg cetaetatte catggtgggg aactgggeta aggttetgat tgtgatgeta 2940
ctetttgeeg gegttgaegg gggaacetat gtgaeagggg ggaegatgge caaaaacace 3000 etegggatta egteetett tteaceeggg teateceaga aaateeaget tgtaaacace 3060 aaeggeaget ggeacateaa eaggaetgee etgaactgea atgaeteeet caacactggg 3120
tteettgetg egetgtteta egtgeacaag tteaacteat etggatgeee agagegeatg 3180
gccagctgca gcccatcga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240 tcacacagct cggaccagag gccttattgt tggcactacg cacccggcc gtgcggtatc 3300 gtacccgcgg cgcaggtgtg tggtccagtg tactgcttca ccccaagccc tgtcgtggtg 3360
gggacgaccg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420
ctgcttctta acaacacgcg gccgccaa ggcaactggt ttggctgtac atggatgaat 3480
agcactgggt toaccaagac gtgcgggggc cccccgtgta acatcggggg gatcggcaat 3540 aaaaccttga cctgccccac ggactgcttc cggaagcacc ccgaggccac ttacaccaag 3600
```

```
tgtggttcgg ggccttggtt gacacccaga tgcttggtcc actacccata caggctttgg 3660
cactacccct gcactgtcaa ctttaccatc ttcaaggtta ggatgtacgt ggggggagtg 3720
gagcacaggc tegaageege atgcaattgg actegaggag agegttgtaa cetggaggae 3780
agggacagat cagagettag ecegetgetg etgtetacaa eggagtggea ggtattgeec 3840
tgttccttca ccaccctacc ggctctgtcc actggtttga tccatctcca tcagaacgtc 3900
gtggacgtac aatacctgta cggtataggg tcggcggttg tctcctttgc aatcaaatgg 3960
gagtatetce tettectett cettettete geggaegege gegtetete etgetteteg 4020
atgatgetge tgatagetea agetgaggee geeetagaga acetggtggt ceteaaegeg 4080
gcatcogtgg coggggcgca tggcattete tectteeteg tgttettetg tgetgeetgg 4140
tacatcaagg gcaggctggt ccctggggcg gcatatgccc tctacggcgt atggccgcta 4200
ctcctgctcc tgctggcgtt accaccacga gcatacgcca tggaccggga gatggcagca 4260
togtgoggag gogoggtttt ogtaggtotg atactottga cottgtoacc goactataag 4320
ctgttcctcg ctaggctcat atggtggtta caatatttta tcaccagggc cgaggcacac 4380
ttgcaagtgt ggatccccc cctcaacgtt cgggggggcc gcgatgccgt catcctcctc 4440
acgtgcgcga tccacccaga gctaatcttt accatcacca aaatcttgct cgccatactc 4500
ggtccactca tggtgctcca ggctggtata accasagtgc cgtacttcgt gcgcgcacac 4560
gggctcattc gtgcatgcat gctggtgcgg aaggttgctg ggggtcatta tgtccaaatg 4620
gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct caccccactg 4680
egggaetggg eccaegeggg ectaegagae ettgeggtgg eagttgagee egtegtette 4740 tetgatatgg agaceaaggt tateacetgg ggggeagaea eegeggegtg tggggaeate 4800
atcitgggcc tgcccgtctc cgcccgcagg gggagggaga tacatctggg accggcagac 4860
ageettgaag ggeaggggtg gegaeteete gegeetatta eggeetaete eeaacagaeg 4920
cgaggectae tiggetgeat cateactage etcacaggee gggacaggaa ccaggtegag 4980
ggggaggtcc aagtggtctc caccgcaaca caatctttcc tggcgacctg cgtcaatggc 5040
gtgtgttgga ctgtctatca tggtgccggc tcaaagaccc ttgccggccc aaagggccca 5100
atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gccccccggg 5160 gcgcgttcct tgacaccatg cacctgcggc agctcggacc tttacttggt cacgaggcat 5220
geogatgica ticoggige ceggeggge gacageaggg ggageetaet etececeagg 5280
eccetetect acttgaaggg ctetteggge ggtecactge tetgececte ggggcacget 5340
gtgggcatct ttcgggctgc cgtgtgcacc cgaggggttg cgaaggcggt ggactttgta 5400
cccgtcgagt ctatggaaac cactatgcgg tccccggtct tcacggacaa ctcgtcccct
                                                                               5460
ccggccgtac cgcagacatt ccaggtggcc catctacacg cccctactgg tagcggcaag 5520
agcactaagg tgccggctgc gtatgcagcc caagggtata aggtgcttgt cctgaacccg 5580
tecgtegeeg ceacectagg ttteggggeg tatatgteta aggeaeatgg tategaeeet 5640
aacatcagaa teggggtaag gaccatcaec aegggtgeec ceatcaegta etecaeetat 5700 ggcaagttte ttgeegaegg tggttgetet gggggegeet atgacatcat aatatgtgat 5760
gagtgccact caactgacte gaccactate etgggcateg geacagteet ggaccaageg 5820
gagacggctg gagcgcgact cgtcgtgctc gccaccgcta cgcctccggg atcggtcacc 5880
gtgccacatc caaacatcga ggaggtggct ctgtccagca ctggagaaat ccccttttat 5940
ggcaaagcca tececatega gaccateaag ggggggagge aceteatttt etgecattee 6000
aagaagaaat gtgatgagct cgccgcgaag ctgtccggcc tcggactcaa tgctgtagca 6060
tattaceggg geettgatgt ateegteata eeaaetageg gagaegteat tgtegtagea 6120
acggacgete taatgacggg etttaceggt gaettegaet eagtgatega etgeaataca 6180
tgtgtcaccc agacagtcga cttcagcctg gacccgacct tcaccattga gacgacgacc 6240
gtgccacaag acgcggtgtc acgctcgcag cggcgaggca ggactggtag gggcaggatg 6300
ggcatttaca ggtttgtgac tecaggagaa eggeeetegg geatgttega tteeteggtt 6360
ctgtgcgagt gctatgacgc gggctgtgct tggtacgagc tcacgcccgc cgagacctca 6420 gttaggttgc gggcttacct aaacacacca gggttgcccg tctgccagga ccatctggag 6480
ttetgggaga gegtetttae aggeeteaee eacatagaeg eecatttett gteecagaet 6540
aagcaggcag gagacaactt cccctacctg gtagcatacc aggctacggt gtgcgccagg 6600 gctcaggctc cacctccatc gtgggaccaa atgtggaagt gtctcatacg gctaaagcct 6660 acgctgcacg ggccaacgcc cctgctgtat aggctgggag ccgttcaaaa cgaggttact 6720
accacacac ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780
acgagcacct gggtgctggt aggcggagtc ctagcagctc tggccgcgta ttgcctgaca 6840
acaggcagcg tggtcattgt gggcaggate atcttgtccg gaaagccggc catcattccc 6900 gacagggaag teetttaccg ggagttcgat gagatggaag agtgcgcctc acacctccct 6960
tacatcgaac agggaatgca gctcgccgaa caattcaaac agaaggcaat cgggttgctg 7020
caaacagcca ccaagcaagc ggaggctgct gctcccgtgg tggaatccaa gtggcggacc 7080 atcgaagcct tctgggcgaa gcatatgtgg aatttcatca gcgggataca atatttagca 7140 ggcttgtcca ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagcctct 7200
atcaccaged egeteaceae ecaacatace etectgttta acateetggg gggatgggtg 7260
geogeocaac tigetectee cagegoiget telegitteg taggegoogg categoigga 7320
geggetgttg geageatagg eettgggaag gtgettgtgg atattttgge aggttatgga 7380 geaggggtgg caggegeet egtggeettt aaggteatga geggegagat geeeteeace 7440
gaggacctgg ttaacctact ccctgctatc ctctcccctg gcgccctagt cgtcggggtc 7500
gtgtgcgcag cgatactgcg tcggcacgtg ggcccagggg agggggctgt gcagtggatg 7560 aaccggctga tagcgttcgc ttcgcggggt aaccacgtct cccccacgca ctatgtgcct 7620
```

```
gagagegaeg etgeageaeg tgteaeteag atecteteta gtettaeeat eacteagetg 7680
ctgaagagge ttcaccagtg gatcaacgag gactgeteca egecatgete eggetegtgg 7740
ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggctccag 7800
tocaagetee tgeegegatt geegggagte ecettettet catgteaacg tgggtacaag
ggagtetgge ggggegaegg catcatgeaa accaectgee catgtgggge acagateaec 7920
ggacatgiga aaaacggtic catgaggate giggggeeta ggacetgiag taacacgigg 7980
catggaacat tececattaa egegtaeace aegggeeeet geaegeeete eeeggegeea 8040
aattatteta gggegetgtg gegggtgget getgaggagt acgtggaggt taegegggtg 8100 ggggatttee actaegtgae gggeatgaee actgaegaeg taaagtgeee gtgteaggtt 8160
coggececcy aattetteae agaagtggat ggggtgeggt tgeacaggta egetecageg 8220
tgcaaacccc tectaeggga ggaggteaca tteetggteg ggeteaatea atacetggtt 8280
gggtcacagc teccatgega gecegaaceg gatgtageag tgetcactte catgetcace 8340 gaccectece acattaegge ggagaegget aagegtagge tggccagggg atetectece 8400
cccttggcca gctcatcagc tagccagctg tctgcgcctt ccttgaaggc aacatgcact 8460
accepteaty actecoegga egetgacete ategaggeea aceteetgtg geggeaggag 8520
atgggcggga acatcacccg cgtggagtca gaaaataagg tagtaatttt ggactctttc 8580
gageegetee aageggagga ggatgagagg gaagtateeg tteeggegga gateetgegg 8640
aggiccagga aattooctog agogatgooc atatgggcac goooggatta caaccotoca 8700
ctgttagagt cctggaagga cccggactac gtccctccag tggtacacgg gtgtccattg 8760 ccgcctgcca aggcccctcc gataccacct tcacggagga agaggacggt tgtcctgtca 8820
gaátotácog tátottotgo éttggoggag etegocadag agacettegg cageteegaa 8880
togtoggoog togacagogg cacggoaacg goototootg accagocoto ogacgaoggo 8940
gacgogggat ccgacgttga gtcgtactcc tccatgcccc cccttgaggg ggagccgggg 9000 gatcccgatc tcagcgacgg gtcttggtct accgtaagcg aggaggctag tgaggacgtc 9060
gretgerger egargrecta cacatggaca ggegecerga reacgecatg egergeggag 9120
gaaaccaagc tgcccatcaa tgcactgagc aactctttgc tccgtcacca caacttggtc 9180
tatgctacaa catctcgcag cgcaagcctg cggcagaaga aggtcacctt tgacagactg 9240
caggteetgg acgaccacta eegggaegtg eteaaggaga tgaaggegaa ggegteeaca 9300
gttaaggeta aacttetate egtggaggaa geetgtaage tgaegeeece acatteggee 9360
agatetaaat ttggetatgg ggeaaaggae gteeggaace tateeageaa ggeegttaae 9420
cacatccgct ccgtgtggaa ggacttgctg gaagacactg agacaccaat tgacaccacc 9480 atcatggcaa aaaatgaggt tttctgcgtc caaccagaga aggggggccg caagccagct 9540
cgccttatcg tattcccaga tttgggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600
gtggtctcca ccctccctca ggccgtgatg ggctcttcat acggattcca atactctcct 9660
ggacageggg tegagtteet ggtgaatgee tggaaagega agaaatgeee tatgggette 9720 geatatgaca eeegetgttt tgaeteaaeg gteaetgaga atgaeateeg tgttgaggag 9780
tcaatctacc aatgttgtga cttggccccc gaagccagac aggccataag gtcgctcaca 9840
gagcggcttt acatcggggg ccccctgact aattctaaag ggcagaactg cggctatcgc 9900
cggtgccgcg cgagcggtgt actgacgacc agctgcggta ataccetcac atgttacttg 9960
aaggeegetg eggeetgteg agetgegaag etecaggaet geaegatget egtatgegga 10020
gacgaccttg togttatctg tgaaagcgcg gggacccaag aggacgaggc gagcctacgg 10080
geetteaegg aggetatgae tagataetet geeceeett gggaeeegee caaaccagaa 10140 tacgaettgg agttgataae atcatgetee tecaatgtgt cagtegegea egatgeatet 10200
ggcaaaaggg tgtactatct cacccgtgac cccaccaccc cccttgcgcg ggctgcgtgg 10260
gagacageta gacacactee agteaattee tggetaggea acateateat gtatgegeee 10320
accttgtggg caaggatgat cetgatgact cattettet ecateettet agetcaggaa 10380
caacttgaaa aagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440
gacctacctc agatcattca acgactccat ggccttagcg cattttcact ccatagttac 10500
tctccaggtg agatcaatag ggtggcttca tgcctcagga aacttggggt accgccttg 10560
cgagtctgga gacatcgggc cagaagtgtc cgcgctaggc tactgtccca gggggggagg 10620
gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680
ccaatcccgg ctgcgtccca gttggattta tccagctggt tcgttgctgg ttacagcggg 10740
ggagacatat atcacageet gtetegtgee egaceceget ggtteatgtg gtgeetacte 10800
ctactttctg taggggtagg catctatcta ctccccaacc gatgaacggg gagctaaaca 10860
ttttttttt tttttttc tcctttttt ttcctcttt tttccttttc tttcctttgg 10980
tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040
agagagtgct gatactggcc tctctgcaga tcaagt
```

```
<211> 8001
<212> DNA
<213> Hepatitis C Virus
<400> 10
gccagecece gattggggge gacactecae catagateae teccetgtga ggaactaetg 60
tetteacgea gaaagegtet agceatggeg ttagtatgag tgtegtgeag cetecaggae 120
ccccctccc gggagagcca tagtggtctg cggaaccggt gagtacaccg gaattgccag 180
gacgaccggg teetteettg gateaacccg etcaatgeet ggagatttgg gegtgeecee 240 gegagactge tagecgagta gtgttgggte gegaaaggee ttgtggtaet geetgatagg 300
gtgettgega gtgeeeggg aggtetegta gaeegtgeae catgageaeg aateetaaac 360
ctcaaagaaa aaccaaacgt aacaccaacg ggcgcgccat gattgaacaa gatggattgc 420
acgeaggtte teeggeeget tgggtggaga ggetattegg etatgaetgg geacaacaga 480 caateggetg etetgatgee geegtgttee ggetgteage geaggggege ceggttettt 540
ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600
egtggetgge caegaeggge gtteettgeg cagetgtget egaegttgte aetgaagegg 660
gaagggactg getgetattg ggcgaagtge eggggeagga teteetgtea teteacettg 720 eteetgeega gaaagtatee ateatggetg atgeaatgeg geggetgeat aegettgate 780
eggetacetg eccattegae caccaagega aacategeat egagegagea egtactegga 840
tggaagcogg tottgtogat caggatgato,tggacgaaga gcatcagggg ctcgcgccag 900
ccgaactgtt cgccaggete aaggegegea tgcccgacgg cgaggatete gtegtgacec 960
atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020
actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080
ttgctgaaga gcttggcggc gaatgggctg accgetteet cgtgetttae ggtategeeg 1140
ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200 cagaccacaa cggttccct ctagcgggat caattccgcc cctctcctc cccccccct 1260
aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320
tecaccatat tgeegtettt tggeaatgtg agggeeegga aacctggeee tgtettettg 1380
acgagcatto ctaggggtot ttoccototo gecaaaggaa tgcaaggtot gttgaatgto 1440
gtgaaggaag cagtteetet ggaagettet tgaagacaaa caaegtetgt agegaeeett 1500
tgcaggcage ggaacceccc acetggcgac aggtgcctct gcggccaaaa gccacgtgta 1560
taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620
gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680
gtaccccatt gtatgggatc tgatctgggg ceteggtgca catgetttae atgtgtttag 1740
togaggttaa aaaacgtota ggccccccqa accacgggga cgtggttttc ctttgaaaaa 1800
cacgataata ccatggcgcc tattacggcc tactcccaac agacgcgagg cctacttggc 1860
tgcátcatca ctagéétéac aggeogggac aggaaccagg tégaggggga ggtccaaggtg 1920 gtetecaccg caacacaate titeotggeg acctgegtea atggegtgtg tiggaetgte 1980
tatcatggtg ccggctcaaa gaccettgcc ggcccaaagg gcccaatcac ccaaatgtac 2040
accaatgtgg accaggacct cgtcggctgg caagcgcccc ccggggcgcg ttccttgaca 2100
ccatgcacct gcggcagete ggacctttac ttggtcacga ggcatgccga tgtcattccg 2160 gtgcgccggc ggggcgacag cagggggage ctactctcc ccaggcccgt ctcctacttg 2220
aagggetett egggeggtee aetgetetge eestegggge aegetgtggg eatetttegg 2280
getgeegtgt geaccegagg ggttgegaag geggtggaet ttgtaceegt egagtetatg 2340 ggaaccacta tgeggteece ggtetteaeg gacaactegt eeetteegge egtacegeag 2400 acatteeagg tggeecatet acaegeeeet actggtageg geaagageae taaggtgeeg 2460
getgegtatg caggecaagg gtataaggtg ettgteetga accegteegt egeegeeace 2520
ctaggtttcg gggcgtatat gtctaaggca catggtatcg accctaacat cagaatcggg 2580
gtaaggacca tcaccacggg tgcccccatc acgtactcca cctatggcaa gtttcttgcc 2640
gacggtggtt getetggggg egeetatgae atcataatat gtgatgagtg ceaetcaact 2700
gactogacca ctatoctggg catoggoaca gtootggaco aagoggagao ggotggagog 2760
egactegteg tgetegeeac egetaegeet eegggategg teacegtgee acatecaaac 2820
ategaggagg tggetetgte eageactgga gaaateeeet tttatggeaa ageeateeee 2880
atcgagacca tcaagggggg gaggcacctc attttctgcc attccaagaa gaaatgtgat 2940
gagetegeeg egaagetigte eggeetegga etcaatgetg tageatatta eeggggeett 3000
gatgtatccg tcataccaac tagcggagac gtcattgtcg tagcaacgga cgctctaatg 3060 acgggettta ccggcgactt cgactcagtg atcgactgca atacatgtgt cacccagaca 3120
gtcgacttca gcctggaccc gaccttcacc attgagacga cgaccgtgcc acaagacgcg 3180
gtgtcacgct cgcagcggcg aggcaggact ggtaggggca ggatgggcat ttacaggttt 3240
gtgactccag gagaacggcc ctcgggcatg ttcgattcct cggttctgtg cgagtgctat 3300
gacgeggget gtgettggta egageteaeg eeegeegaga eeteagttag gttgeggget 3360
tacctaaaca caccagggtt gcccgtctgc caggaccatc tggagttctg ggagagcgtc 3420
tttacaggee teaceacat agaegeeeat ttettgteee agaetaagea ggeaggagae 3480
aacttocoot acctggtago afaccaggot acggtgtgcg ccagggotca ggotocacot 3540
ccatcgtggg accaaatgtg gaagtgtctc atacggctaa agcctacgct gcacgggcca 3600
```

```
acgcccctgc tgtataggct gggagccgtt caaaacgagg ttactaccac acaccccata 3660
accaaataca tcatggcatg catgtcggct gacctggagg tcgtcacgag cacctgggtg 3720
 etggtaggeg gagteetage agetetggee gegtattgee tgacaacagg cagegtggte 3780
attgtgggca ggatcatett gtocggaaag coggccatca ttcccgacag ggaagtcctt 3840
taccgggagt tcgatgagat ggaagagtgc gcctcacacc tcccttacat cgaacaggga 3900
atgcagctcg ccgaacaatt caaacagaag gcaatcgggt tgctgcaaac agccaccaag 3960
caageggagg etgetgetee egtggtggaa tecaagtgge ggaceatega ageettetgg 4020
gcgaagcata tgtggaattt catcagcggg atacaatatt tagcaggctt gtccactctg 4080
cotggcaacc ocgogatage atcactgatg geatteacag cototateac cageocgete 4140
accacceaac ataccetect gtttaacate etggggggat gggtggeege ceaacttget 4200
cctcccagcg ctgcttctgc tttcgtaggc gccggcatcg ctggagcggc tgttggcagc 4260 ataggccttg ggaaggtgct tgtggatatt ttggcaggt atggagcagg ggtggcaggc 4320
gegetegtgg cetttaaggt catgagegge gagttgeeet eeacegagga cetggttaac 4380
etactecetg etatectete ecetggegee etagtegteg gggtegtgtg egeagegata 4440
ctgcgtcggc acgtgggccc aggggaggg gctgtgcagt ggatgaaccg gctgatagcg 4500 ttcgcttcgc ggggtaacca cgtctcccc acgcactatg tgcctgagag cgacgctgca 4560
gcacgagtea etcagateet etetagtett accateacte agetgetgaa gaggetteae 4620
cagtggatca acgaggactg ctccacgeca tgeteegget egtggetaag agatgtttgg 4680
gattggatat geaeggtgtt gactgattte aagacetgge tecagteeaa geteetgeeg 4740
cgattgccgg gagtcccctt cttctcatgt caacgtgggt acaagggagt ctggcggggc 4800
gacggcatca tgcaaaccac ctgcccatgt ggagcacaga tcaccggaca tgtgaaaaac 4860
ggttccatga ggatcgtggg gcctaggacc tgtagtaaca cgtggcatgg aacattcccc 4920
attaacgcgt acaccacggg cccctgcacg ccctccccgg cgccaaatta ttctagggcg 4980
ctgtggcggg taggtgctga ggagtacgtg gaggttacgc gggtggggga tttccactac 5040
gtgacgggca tgaccactga caacgtaaag tgcccgtgtc aggttccggc ccccgaattc 5100
ttcacagaag tggatggggt geggttgcac aggtacgete cagegtgcaa acceetecta 5160
cgggaggagg tcacattcct ggtcgggctc aatcaatacc tggttgggtc acagctccca 5220
tgcgagcctg aacoggatgt agcagtgctc acttccatgc tcaccgaccc ctcccacatt 5280
acggcggaga cggctaagcg taggctggcc aggggatctc ccccccctt ggccagctca 5340
teagetagee agetgtetge geetteettg aaggeaacat geactaceeg teatgaetee 5400 eeggaegetg aceteatega ggeeaacete etgtggegge aggagatggg egggaacate 5460
accegegtgg agteagaaaa taaggtagta attttggact etttegagee geteeaageg 5520
gaggaggatg agaggggagt atccgttccg gcggagatcc tgcggaggtc caggaaattc 5580
cetegagega tgeccatatg ggeaegeeeg gattacaace etecaetgtt agagteetgg 5640 aaggaeeegg actaegteee teeagtggta caegggtgte cattgeegee tgecaaggee 5700
cctccgatac caccttcacg gaggaagagg acggttgtcc tgtcagaatc taccgtgtct 5760
totgoottgg oggagotogo cacagagaco ttoggoagot cogaatogto ggoogtogae 5820
ageggeaegg caaeggeete teetgaeeag eesteegaeg aeggegaege gggateegae 5880
gttgagtcgt actcctccat gcccccctt gagggggagc cgggggatcc cgatctcagc 5940
gacgggtctt ggtctaccgt aagcgaggag gctagtgagg acgtcgtctg ctgctcgatg 6000
tectacacat ggacaggege cetgateacg ceatgegetg eggaggaaac caagetgeee 6060
atcaatgcac tgagcaacte tttgeteegt caccacaact tggtetatge tacaacatet 6120
cgcagcgcaa acctgcggca gaagaaggtc acctttgaca gactgcaggt cctggacgac 6180
cactaccggg acgtgctcaa ggagatgaag gcgaaggcgt ccacagttaa ggctaaactt 6240
ctatccgtgg aggaagcctg taagctgacg ccccacatt cggccagatc taaatttggc 6300
tatggggcaa aggacgtccg gaacctatcc agcaaggccg ttaaccacat ccgctccgtg 6360 tggaaggact tgctggaaga cactgagaca ccaattgaca ccaccatcat ggcaaaaaat 6420
gaggttttct gcgtccaacc agagaagggg ggccgcaagc cagctcgcct tatcgtattc 6480
ccagatttgg gggttcgtgt gtgcgagaaa atggcccttt acgatgtggt ctccaccctc 6540
ceteaggeeg tgatgggete tteataegga tteeaataet eteetggaea gegggtegag 6600
tteetggtga atgeetggaa agegaagaaa tgeeetatgg gettegeata tgaeaeeege 6660
tgttttgact caacggtcac tgagaatgac atccgtgttg aggagtcaat ctaccaatgt 6720
tgtgacttgg cccccgaagc cagacaggcc ataaggtcgc tcacagagcg gctttacatc 6780
gggggccccc tgactaattc taaagggcag aactgcggct atcgccggtg ccgcgcgagc 6840
ggtgtactga cgaccagctg cggtaatacc ctcacatgtt acttgaaggc cgctgcggcc 6900
tgtcgagctg cgaagctcca ggactgcacg atgctcgtat gcggagacga ccttgtcgtt 6960
atotgtgaaa gogoggggac ccaagaggac gaggogagoc tacggggcott cacggaggot 7020
atgactagat actotgococ coctggggac cogocoaaac cagaatacga cttggagttg 7080 ataacatcat geteeteeaa tgtgteagte gegeacgatg catetggeaa aagggtgtac 7140
tateteacce gtgaceceae caceceett gegegggetg egtgggagae agetagacae 7200
actocagtea attectgget aggeaacate ateatgtatg egeceacett gtgggeaagg 7260
atgatectga tgactcattt citetecate ettetagete aggaacaact tgaaaaagee 7320
ctagattgtc agatctacgg ggcctgttac tccattgagc cacttgacct acctcagatc 7380
atteaacgae tecatggeet tagegeattt teactecata gttactetee aggtgagate 7440
aatagggtgg cttcatgcct caggaaactt ggggtaccgc ccttgcgagt ctggagacat 7500
egggecagaa gtgteegege taggetactg teecaggggg ggagggetge caettgtgge 7560 aagtacetet teaactggge agtaaggace aageteaaac teacteeaat eeeggetgeg 7620
```

tcccagttgg	atttatccag	ctggttcgtt	gctggttaca	gcgggggaga	catatatcac	7680
agcctgtctc	gtgcccgacc	ccgctggttc	atgtggtgcc	tactcctact	ttctgtaggg	7740
gtaggcatct	atctactccc	caaccgatga	acggggagct	aaacactcca	ggccaatagg	7800
ccatcctgtt	tttttccctt	tttttttc	tttttttt	tttttttt	tttttttt	7860
ttttctcctt	tttttttcct	cttttttcc	ttttctttcc	tttggtggct	ccatcttage	7920
cctagtcacg	gctagctgtg	aaaggtccgt	gagccgcttg	actgcagaga	gtgctgatac	7980
tggcctctct	gcagatcaag	t			<del>-</del>	8001

```
<210> 11
<211> 11076
<212> DNA
<213> Hepatitis C Virus
```

<400> 11 gccagccccc gattgggggc gacactccac catagatcac tcccctgtga ggaactactg 60 tetteaegea gaaagegtet ageeatggeg ttagtatgag tgtegtgeag ceteeaggae 120 eccectece gggagageea tagtggtetg eggaaceggt gagtacaceg gaattgeeag 180 gacgaccggg teettettg gateaacccg etcaatgeet ggagatttgg gegtgeecce 240 gcgagactgc tagccgagta gtgttgggtc gcgaaaggcc ttgtgggtact gcctgatagg 300 gtgcttgcga gtgccccggg aggtctcgta gaccgtgcac catgagcacg aatcctaaac 360 cicaaagaaa aaccaaacgi aacaccaacg ggcgcgccat gatigaacaa gatggattgc 420 acgcaggttc teeggeeget tgggtggaga ggctattegg ctatgaetgg gcacaacaga 480 caateggetg etetgatgee geogtgttee ggetgteage geaggggege ceggttettt 540 ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca gcgcggctat 600 cytgyctgyc cacgacgygc yttocttycy cagetytyct cyacyttyte actgaagegy 660 gaagggacty yetgetatty gycgaaytyc cygygcagga teteetytea teteacetty 720 ctcctgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat acgcttgatc 780 eggetacetg eccattegae caccaagega aacategeat egagegagea egtactegga 840 tggaageegg tettgtegat caggatgate tggacgaaga geatcagggg etegegeeag 900 cegaactgtt egecaggete aaggegegea tgeeegaegg egaggatete gtegtgaeee 960 atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct ggattcatcg 1020 actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct acccgtgata 1080 ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg 1140 ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc tgagtttaaa 1200 cagaccacaa cggtttccct ctagcgggat caattccgcc cctctccctc cccccccct 1260 aacgttactg gccgaagccg cttggaataa ggccggtgtg cgtttgtcta tatgttattt 1320 tccaccatat tgccgtcttt tggcaatgtg agggcccgga aacctggccc tgtcttcttg 1380 acgageatte ctaggggtet tteceetete gecaaaggaa tgeaaggtet gttgaatgte 1440 gtgaaggaag cagttcctct ggaagcttct tgaagacaaa caacgtctgt agcgaccctt 1500 tgcaggcage ggaacccccc acctggcgac aggtgcctct gcggccaaaa gccacgtgta 1560 taagatacac ctgcaaaggc ggcacaaccc cagtgccacg ttgtgagttg gatagttgtg 1620 gaaagagtca aatggctctc ctcaagcgta ttcaacaagg ggctgaagga tgcccagaag 1680 gtaccccatt gtatgggatc tgatctgggg cctcggtgca catgctttac atgtgtttag 1740 tegaggttaa aaaaegteta ggeeeeega aecaegggga egtggtttte etttgaaaaa 1800 cacgataata ccatgggcac gaatcctaaa cctcaaagaa aaaccaaacg taacaccaac 1860 cgccgcccac aggacgicaa gitcccgggc ggtggtcaga tcgtcggtgg agittacctg 1920 ttgccgcgca ggggccccag gttgggtgtg cgcgcgacta ggaagacttc cgagcggtcg 1980 caacetegtg gaaggegaca acetateece aaggetegee ageeegaggg tagggeetgg 2040 gctcagcccg ggtacccctg gcccctctat ggcaatgagg gcttggggtg ggcaggatgg 2100 ctcctgtcac cccgtggctc tcggcctagt tggggcccca cggacccccg gcgtaggtcg 2160 egeaatttgg gtaaggteat egataceete aegtgegget tegeegatet eatggggtae 2220 attecgeteg teggegeece cetaggggge getgecaggg ceetggegea tggegteegg 2280 gttctggagg acggcgtgaa ctatgcaaca gggaatctgc ccggttgctc cttttctatc 2340 ttccttttgg ctttgctgtc ctgtttgacc atcccagctt ccgcttatga agtgcgcaac 2400 gtatccggag tgtaccatgt cacgaacgac tgctccaacg caagcattgt gtatgaggca 2460 geggacatga teatgeatae eccegggtge gtgeeetgeg ttegggagaa caacteetee 2520 egetgetggg tagegeteae teccaegete geggeeagga acgetagegt ecceaetacg 2580 acgataegae gecatgtega titgetegti ggggeggetg etetetgete egetatgtae 2640 gtgggagate tetgeggate tgtttteete gtegeecage tgtteacett etegeetege 2700 eggeacgaga cagtacagga etgeaattge teaatatate eeggeeacgt gacaggteae 2760 egtatggett gggatatgat gatgaactgg teacetacag cageectagt ggtategeag 2820 ttactcogga tcccacaago tgtcgtggat atggtggcgg gggcccattg gggagtccta 2880 gegggeettg cetactatte catggtgggg aactgggeta aggttetgat tgtgatgeta 2940 ctctttgccg gcgttgacgg gggaacctat gtgacagggg ggacgatggc caaaaacacc 3000 ctcgggatta cgtccctctt ttcacccggg tcatcccaga aaatccagct tgtaaacacc 3060 aacggcagct ggcacatcaa caggactgcc ctgaactgca atgactccct caacactggg 3120 tteettgetg egetgtteta egtgeacaag tteaacteat etggatgeec agagegeatg 3180 gccagctgca gccccatcga cgcgttcgct caggggtggg ggcccatcac ttacaatgag 3240 tcacacaget eggaceagag geettatigt tggcactacg cacceeggee gtgeggtate 3300 gtaceegegg egeaggigtg tggtceagig tactgettea ecceaagee tgtegtggtg 3360 gggacgaccg accggttcgg cgtccctacg tacagttggg gggagaatga gacggacgtg 3420 ctgcttctta acaacacgcg gccgccgcaa ggcaactggt ttggctgtac atggatgaat 3480 ageactgggt teaccaagae gtgegggge ecceegtgta acateggggg gateggeaat 3540 aaaaeettga eetgeeeca ggaetgette eggaageaee eegaggeeae ttacaceaag 3600

```
tgtggttcgg ggccttggtt gacacccaga tgcttggtcc actacccata caggctttgg 3660
 cactacccct gcactgtcaa ctttaccatc ttcaaggtta ggatgtacgt ggggggagtg 3720
 gagcacagge tegaageege atgeaattgg actegaggag agegttgtaa eetggaggae 3780
 agggacagat cagagettag ecegetgetg etgtetacaa eggagtggea ggtattgeee 3840
 tgitcettca ccaccetace ggetetgice actggittga tecateteca teagaaegte 3900
 gtggacgtac aatacctgta cggtataggg tcggcggttg tctcctttgc aatcaaatgg 3960
 gagtatgtee tgttgetett cettettetg geggaegege gegtetgtge etgettgtgg 4020
 atgatgetge tgatagetca agetgaggee geeetagaga acetggtggt ceteaacgeg 4080
 geatecgtgg eegggegea tggeattete teetteeteg tgttettetg tgetgeetgg 4140 tacateaagg geaggetggt eeetggggeg geatatgee tetaeggegt atggeegeta 4200
 ctcctgctcc tgctggcgtt accaccacga gcatacgcca tggaccggga gatggcagca 4260
 tegtgeggag gegeggtttt egtaggtetg atactettga cettgteace geactataag 4320
 ctgttcctcg ctaggctcat atggtggtta caatatttta tcaccagggc cgaggcacac 4380
 ttgcaagtgt ggatccccc cctcaacgtt cggggggcc gcgatgccgt catcctcctc 4440
 acqtqcqcqa tccacccaga gctaatcttt accatcacca aaatcttqct cgccatactc 4500
 ggtccactca tggtgctcca ggctggtata accaaagtgc cgtacttcgt gcgcgcacac 4560
 gggctcattc gtgcatgcat gctggtgcgg aaggttgctg ggggtcatta tgtccaaatg 4620 gctctcatga agttggccgc actgacaggt acgtacgttt atgaccatct caccccactg 4680
 cgggactggg cccacgeggg cctacgagac cttgeggtgg cagttgagec egtegtette 4740
 tetgatatgg agaceaaggt tateacetgg ggggeagaea eegeggegtg tggggaeate 4800
atcttgggcc tgcccgtctc cgcccgcagg gggagggaga tacatctggg accggcagac 4860 agccttgaag ggcaggggtg gcgactcctc gcgcctatta cggcctactc ccaacagacg 4920
 cgaggeetae ttggetgeat cateaetage eteacaggee gggaeaggaa ccaggtegag 4980
 ggggaggtee aagtggtete cacegeaaca caatetttee tggegaeetg egteaatgge 5040
 gtgtgttgga ctgtctatca tggtgccggc tcaaagaccc ttgccggccc aaagggccca 5100
atcacccaaa tgtacaccaa tgtggaccag gacctcgtcg gctggcaagc gccccccggg 5160
gegegtteet tgacaccatg cacctgegge ageteggace tttacttggt cacgaggeat 5220
gccgatgtca ttccggtgcg ccggcggggc gacagcaggg ggagcctact ctccccagg 5280 cccgtctcct acttgaaggg ctcttcggc ggtccactgc tctgcccctc ggggcacgct 5340 gtgggcatct ttcgggctgc cgtgtgcacc cgaggggttg cgaaggcggt ggactttgta 5400
eccgtegagt ctatgggaac cactatgegg tecceggtet teaeggacaa etegteeet 5460
ccggccgtac cgcagacatt ccaggtggcc catctacacg cccctactgg tagcggcaag 5520
agcactaagg tgccggctgc gtatgcaggc caagggtata aggtgcttgt cctgaacccg 5580
teegtegeeg ceacectagg ttteggggeg tatatgteta aggeacatgg tategaceet 5640
aacatcagaa teggggtaag gaccateace aegggtgeee eeateaegta etecaeetat 5700
ggcaagttte ttgccgacgg tggttgctct gggggcgcct atgacatcat aatatgtgat 5760
gagtgccact caactgacte gaccactate etgggcateg gcacagteet ggaccaageg 5820
gagacggctg gagcgcgact cgtcgtgctc gccaccgcta cgcctccggg atcggtcacc 5880
gtgccacato caaacatoga ggaggtggot otgtocagoa otggagaaat coccitttat 5940
ggcaaagcca teeccatega gaccateaag ggggggagge aceteatttt etgecattee 6000
aagaagaaat gtgatgaget egeeggaag etgteeggee teggaeteaa tgetgtagea 6060 tattaceggg geettgatgt ateegteata eeaactageg gagaegteat tgtegtagea 6120
acggacgete taatgacggg etttacegge gaettegaet eagtgatega etgeaataca 6180
tgtgtcaccc agacagtcga cttcagectg gacccgacct tcaccattga gacgacgacc 6240
gtgccacaag acgcggtgtc acgctcgcag cggcgaggca ggactggtag gggcaggatg 6300
ggcatttaca ggtttgtgac tocaggagaa cggccctcgg gcatgttcga ttcctcggtt 6360
ctgtgcgagt gctatgacgc gggctgtgct"tggtacgagc tcacgcccgc cgagacctca 6420
gttaggttgc gggcttacct aaacacacca gggttgcccg tctgccagga ccatctggag 6480
ttctgggaga gcgtctttac aggcctcacc cacatagacg cccatttctt gtcccagact 6540 aagcaggcag gagacaactt cccctacctg gtagcatacc aggctacggt gtgcgccagg 6600
getcaggete caectecate gtgggaccaa atgtggaagt gtetcatacg getaaageet 6660
acgotgoacg ggocaacgoo cotgotgtat aggotgggag cogttoaaaa ogaggttact 6720
accacacac ccataaccaa atacatcatg gcatgcatgt cggctgacct ggaggtcgtc 6780 acgagcacct gggtgctggt aggcggagtc ctagcagctc tggccgcgta ttgcctgaca 6840
acaggcagcg tggtcattgt gggcaggato atottgtcog gaaagcoggo catcattoco 6900
gacagggaag teetttaceg ggagttegat gagatggaag agtgegeete acaceteeet 6960
tacatogaac agggaatgca gotogoogaa caattoaaac agaaggcaat ogggttgotg 7020 caaacagcca ocaagcaago ggaggotgot gotocogtgg tggaatccaa gtggoggaco 7080
atogaagoot totgggogaa goatatgtgg aatttoatoa gogggataca atatttagoa 7140
ggcttgtcca ctctgcctgg caaccccgcg atagcatcac tgatggcatt cacagcctct 7200
atcaccagee egeteaceae ecaacatace etectottta acateetogo oggatogoto 7260
geogeocaac tigeteetee cagegeiget tetgetiteg taggegeegg categeigga 7320
geggetgttg geageatagg eettgggaag gtgettgtgg atattttgge aggttatgga 7380
gcaggggtgg caggcgcgct cgtggccttt aaggtcatga gcggcgagtt gccctccacc 7440
gaggacotgg ttaacotact coetgetate eteteceetg gegecetagt egteggggte 7500 gtgtgegeag egatactgeg teggeacgtg ggeceagggg agggggetgt geagtggatg 7560
aaccggctga tagcgttcgc ttcgcggggt aaccacgtct cccccacgca ctatqtqcct 7620
```

```
gagagegaeg etgeageaeg agteacteag atecteteta gtettaceat cacteagetg 7680
 ctgaagaggc ttcaccagtg gatcaacgag gactgctcca cgccatgctc cggctcgtgg 7740
 ctaagagatg tttgggattg gatatgcacg gtgttgactg atttcaagac ctggctccag 7800
 tecaagetee tgeegegatt geegggagte ecettettet catgteaacg tgggtacaag 7860
 ggagtctggc ggggcgacgg catcatgcaa accacctgcc catgtggagc acagatcacc 7920
 ggacatgiga aaaacggtic catgaggato gtggggccta ggaccigtag taacacgtgg 7980
 catggaacat tececattaa egegtaeaee aegggeeeet geaegeeete eeeggegeea 8040
 aattattota gggcgctgtg gcgggtaggt gctgaggagt acgtggaggt tacgcgggtg 8100
 ggggatttee actaegtgae gggeatgaee actgaeaacg taaagtgeee gtgteaggtt 8160
 coggcccccg aattetteac agaagtggat ggggtgeggt tgcacaggta cgctccageg 8220
 tgcaaacccc tcctacggga ggaggtcaca ttcctggtcg ggctcaatca atacctggtt 8280
 gggtcacage teccatgega geetgaaceg gatgtageag tgetcactte catgetcace 8340
 gacccctccc acattacggc ggagacggct aagcgtaggc tggccagggg atctccccc 8400
 ccettggcca geteateage tagecagetg tetgegeett cettgaagge aacatgeact 8460
 acceptcatg acteccegga egetgacete ategaggeca acetectgtg geggeaggag 8520
 atgggcggga acatcacccg cgtggagtca gaaaataagg tagtaatttt ggactctttc 8580
 gageegetee aageggagga ggatgagagg ggagtateeg tteeggegga gateetgegg 8640 aggteeagga aatteeeteg agegatgeee atatgggeae geeeggatta caaceeteea 8700
 ctgttagagt cctggaagga cccggactac gtccctccag tggtacacgg gtgtccattg 8760
 ccgcctgcca aggcccctcc gataccacct tcacggagga agaggacggt tgtcctgtca 8820
 gaatetaccg tgtettetge ettggeggag etegecaeag agaeettegg eageteegaa 8880
 togtoggoog togacagogg cacggoaacg goototootg accagoooto cgacgacggo 8940
 gacgegggat cegacgttga gtegtactee tecatgeece ecettgaggg ggageegggg 9000
 gatecegate teagegaegg gtettggtet accgtaageg aggaggetag tgaggaegte 9060
gtetgetget egatgteeta cacatggaca ggegeeetga teaegeeatg egetgeggag 9120 gaaaccaage tgeecatcaa tgeaetgage aactetttge teegteacca caacteggte 9180
 tatgctacaa catctcgcag cgcaaacctg cggcagaaga aggtcacctt tgacagactg 9240
 caggiccigg acgaccacta ccgggacgig cicaaggaga igaaggcgaa ggcgiccaca 9300
gttaaggcta aacttctatc cgtggaggaa gcctgtaagc tgacgccccc acattcggcc 9360
 agatetaaat ttggetatgg ggeaaaggae gteeggaace tateeageaa ggeegttaae 9420
cacatcogot cogtgtggaa ggacttgctg gaagacactg agacaccaat tgacaccacc 9480
atcatggcaa aaaatgaggt tttctgcgtc caaccagaga aggggggccg caagccagct 9540
cgccttatcg tattcccaga tttgggggtt cgtgtgtgcg agaaaatggc cctttacgat 9600 gtggtctcca ccctccctca ggccgtgatg ggctcttcat acggattcca atactctcct 9660
ggacagcggg tcgagttcct ggtgaatgcc tggaaagcga agaaatgccc tatgggcttc 9720
gcatatgaca cccgctgttt tgactcaacg gtcactgaga atgacatccg tgttgaggag 9780
tcaatctacc aatgttgtga cttggccccc gaagccagac aggccataag gtcgctcaca 9840
gageggettt acateggggg ecceetgact aattetaaag ggeagaactg eggetatege 9900
eggtgeegeg egageggtgt actgaegace agetgeggta atacceteae atgttacttq 9960
aaggeegetg eggeetgteg agetgegaag etceaggaet geacgatget egtatgegga 10020
gacgaccttg tcgttatctg tgaaagcgcg gggacccaag aggacgaggc gagcctacgg 10080 gccttcacgg aggctatgac tagatactct gcccccctg gggacccgcc caaaccagaa 10140 tacgacttgg agttgataac atcatgctcc tccaatgtgt cagtcgcgca cgatgcatct 10200
ggcaaaaggg tgtactatet cacccgtgac cccaccaccc cccttgcgcg ggctgcgtgg 10260
gagacageta gacacaetee agteaattee tggetaggea acateateat gtatgegeee 10320
accttgtggg caaggatgat cctgatgact catttcttct ccatccttct agctcaggaa 10380
caacttgaaa aagccctaga ttgtcagatc tacggggcct gttactccat tgagccactt 10440
gacctacete agateattea acgaetecat ggeettageg cattiteact ceatagitae 10500
totocaggig agaicaatag ggiggotica tgootoagga aactiggggt accgcccitg 10560
cgagtctgga gacatcgggc cagaagtgtc cgcgctaggc tactgtccca gggggggagg 10620 gctgccactt gtggcaagta cctcttcaac tgggcagtaa ggaccaagct caaactcact 10680
ccaatcccgg ctgcgtccca gttggattta tccagctggt tcgttgctgg ttacagcggg 10740
ggagacatat atcacageet gretegtgee egaceeeget ggtteatgtg gtgeetaete 10800
ttttttttt tttttttc tcctttttt ttcctcttt tttccttttc tttcctttgg 10980
tggctccatc ttagccctag tcacggctag ctgtgaaagg tccgtgagcc gcttgactgc 11040 agagagtgct gatactggcc tctctgcaga tcaagt 11076
```

#### CLAIMS

- A hepatitis C virus (HCV) cell culture system, which comprises mainly eukaryotic cells containing 5 transfected HCV specific genetic material, characterized in that, the eukaryotic cells are human hepatoma cells and the transfected HCV specific genetic material is a HCV-RNA construct, which comprises the HCV specific segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 10 3' NTR as well as an additional marker gene for selection (selection gene).
- A cell culture system according to Claim 1, characterized in that, the hepatoma cells are derived from a usual hepatoma cell line.
- 3. A cell culture system according to Claim 1, characterized in that, the hepatoma cells are extracted from a primary hepatoma cell culture.
- 4. A HCV-RNA construct, characterized in that, 25 it comprises the HCV specific RNA segments 5' NTR, NS3, NS4A, NS4B, NS5A, NS5B, and 3' NTR as well as an additional marker gene for selection (selection gene).
- 5. A HCV-RNA construct according to Claim 4,
  characterized in that,
  it comprises one of the nucleotide sequences
  illustrated in the sequence protocols SEQ ID NO: 1 to
  SEQ ID NO: 11.
- 35 6. A HCV-RNA construct according to Claim 4, characterized in that,

.10

the	3' NTR has a nucleotide sequence, which has been
	ected from the group of nucleotide sequences (a) to
	listed in the following:
(a)	ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
	GCTTTTTTTTTTTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC CAGACTGGAGT
(b)	ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
	GTCTTTTTTTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCCCA GACTGGAGC
(c)	ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
	CTTTTTTTTTTTTTTTTTTGAGAGAGAGTCTCACTCTGTTGCCCA GACTGCAGC
(d)	ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
	AGTCTTTTTTTTTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGC CCAGACTGGAGT
(e)	ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
٠	GTCTTTTTTTTTTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGC CCAGACTGGAGT
(f)	ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
	${\tt GTCTTTTTTTTTTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTG}$
	CCCAGACTGGAGT
(g)	ACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTTTTTT
	${\tt GTCTTTTTTTTTTTTTTTTTTGAGAGAGAGAGTCTCACTCTGTTGCC}$
	CAGACTGGAGT

TGCCCAGACTGGAGT

٠.

- 5 7. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that, the marker gene for selection is a resistance gene and preferably an antibiotic resistance gene.
- 10 8. A HCV-RNA construct according to one of the Claims 4 to 6, characterized in that, the marker gene for selection is a neomycin phosphotransferase gene.
- 15 9. A HCV-RNA construct according to one of the Claims 4 to 8, characterized in that, the marker gene for selection is integrated in the HCV-RNA downstream from the 5' NTR.
- 20 10. A HCV-RNA construct according to one of the Claims 4 to 83, characterized in that, the marker gene for selection is linked to the HCV-RNA via a ribozyme or a recognition sequence for a ribozyme.

25

- 11. A HCV-RNA construct according to one of the Claims 4 to 10, characterized in that, it has an integrated reporter gene.
- 12. 30 Α HCV-RNA construct according to Claim 11, characterized in that, the reporter gene is a gene from the group of the luciferase genes, the CAT gene (chloramphenicol acetyl transferase gene), the lacZ gene (beta galactosidase 35 gene), the GFP gene (green fluorescence protein gene), the GUS gene (glucuronidase gene) and the SEAP gene (secreted alkaline phosphatase gene).

- 13. A HCV-RNA construct according to one of the Claims 4 to 11, characterized in that, their replication has an influence on the expression of a (cellular) surrogate marker gene.
- 14. A HCV-RNA construct according to one of the Claims 11 to 13, characterized in that, the resistance gene is cloned into the open reading frame of the HCV-RNA in such a way that it will only be transferred to an active form after proteolytic processing.

- 15. A HCV-RNA construct according to one of the Claims 11 to 14, characterized in that, the reporter gene and the marker gene for selection are arranged in the construct in such a way, that they are expressed as a fusion protein.
- 20 16. A cell culture system according to one of the Claims 1 to 3, characterized in that, the HCV-RNA construct is a construct according to at least one of the Claims 4 to 15.
- 25 17. A cell culture system according to Claim 1, characterized in that, the cells containing the HCV-RNA construct are lodged at the DSMZ, Braunschweig, Germany under the deposit number DSM ACC2394 (laboratory name HuBl 9-13).
- 18. A use of a cell culture system according to one of the Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct according to one of the Claims 4 to 15 for the production and/or evaluation and/or testing of treatments and/or diagnostics for the treatment of HCV infections in particular.

- 19. A use of a cell culture system according to one of the Claims 1 to 3 or 16 to 17 and/or a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a vaccine against HCV infections.
- 20. A use a HCV-RNA construct according to one of the Claims 4 to 15 for the production of a liver cell specific vector in gene therapy.

15

25

30

35

- 10 21. A HCV-RNA construct according to one of the Claims 4 to 15, characterized in that, it has an integrated foreign gene and can be used to transfect this foreign gene in a target cell, which can be used to express this foreign gene.
- 22. A process for the production of cell culture adapted mutants of a HCV-RNA construct according to one of the claims 4 to 15, in which the mutants have increased replication efficiency compared to the HCV-RNA construct,
  - characterized in that, a cell culture system according to claim 1, in which the transfected HCV specific genetic material is a HCV-RNA construct with a selection gene according to one of the claims 4 to 15, is cultivated on/in the selection medium corresponding to the selection gene, that the cell clones created are collected, and that the HCV-RNA constructs or parts thereof are isolated
  - 23. A process according to claim 22, characterized in that,
    the isolated HCV-RNA constructs are passaged at least one more time. That is, they are transfected in cells of a cell culture system according to claim 1. The cell culture system of claim 1 thereby obtained, in

which the transfected HCV specific genetic material is

from these cell clones.

the isolated HCV-RNA construct with a selection gene, is cultivated on/in the selection medium corresponding to the selection gene, the cell clones created are collected, and the HCV-RNA constructs are isolated from these cell clones.

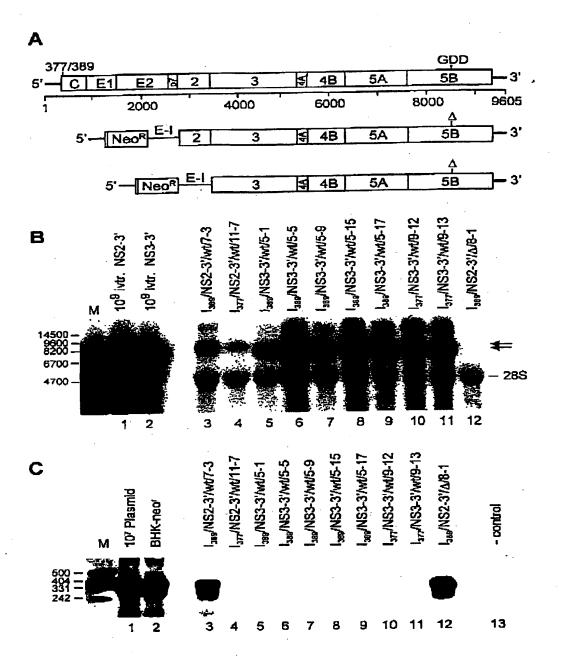
- 24. A process for the production of mutants of a HCV fulllength genome or of a HCV subgenome or any HCV construct with increased replication efficiency in 10 comparison to the original HCV full-length genome or subgenome or HCV-RNA construct, characterized in that, a cell culture adapted mutant of a HCV-RNA construct is produced and isolated by a process according to 15 claim 22 or 23. the nucleotide and amino acid sequence of mutants is determined and the type, number positions of the nucleotide and amino acid mutations is determined by comparison with the nucleotide and 20 amino acid sequence of the original HCV-RNA construct, and these mutations are introduced in an (isolated) HCV full-length genome or a HCV subgenome or any HCV-RNA construct, either by site-directed mutagenesis or by exchange of fragments containing the relevant
  - 25. A cell culture adapted HCV-RNA construct with high replication efficiency, characterized in that,
- it is derivable from a HCV-RNA construct according to one of the claims 4 to 15 through nucleotide and/or amino acid mutations and that it is obtainable with a process according to one of the claims 22 to 24.

25

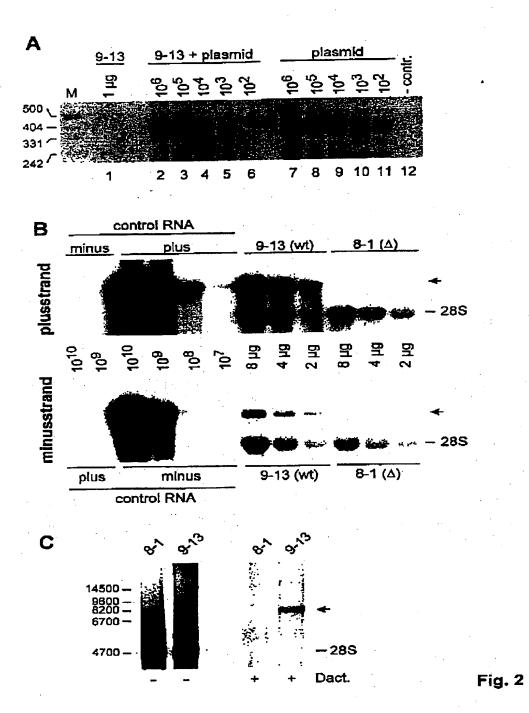
mutations

35 26. A cell culture adapted HCV-RNA construct according to claim 25, characterized in that, it carries one or several of the following amino acid exchanges: 1283 arg -> gly and/or 1383 glu -> ala and/or 1577 lys -> arg and/or 1609 lys -> glu and/or 1936 pro -> ser and/or 2163 glu -> gly and/or 2330 lys -> glu and/or 2442 ile -> val.

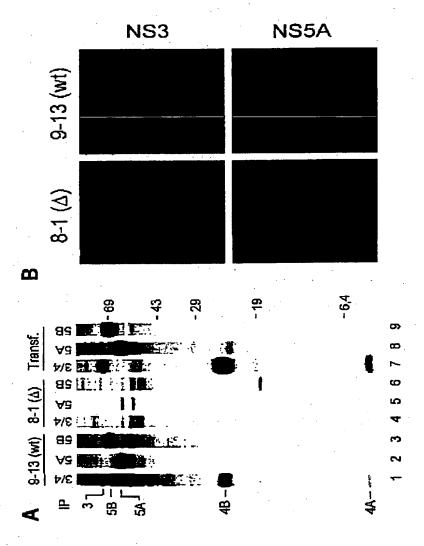
- 27. A cell culture adapted HCV-RNA construct according to claim 25 or 26, characterized in that,
- it displays one or several of the nucleotide and/or amino acid exchanges shown in Table 3, where Table 3 is an integral part of this claim.
- 28. Cell culture adapted mutant of a HCV-RNA construct or of a HCV full-length genome with increased replication efficiency in comparison to the original HCV-RNA construct or original HCV full-length genome, characterized in that,
- it is obtainable with a process in which the number and type of mutations in a cell culture adapted HCV-RNA construct according to claim 24 are determined through sequence analysis and sequence comparison, and these mutations are introduced in a HCV-RNA construct, particularly in a HCV-RNA construct according to one
- of the claims 4 to 15, or in an (isolated) HCV-RNA full-length genome, either by targeted mutagenesis or by exchange of sequence sections containing the relevant mutations.
- 30 29. Hepatitis C virus particles or virus-like particles, characterized in that, they are obtainable by a process according to one of the claims 22 - 24.
- 35 30. Cells infected with Hepatitis C virus particles or virus-like particles, according to claim 29.



Flg. 1



Borden Ladner Gervais LLP | s.r.l.



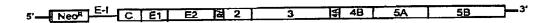


Fig. 4

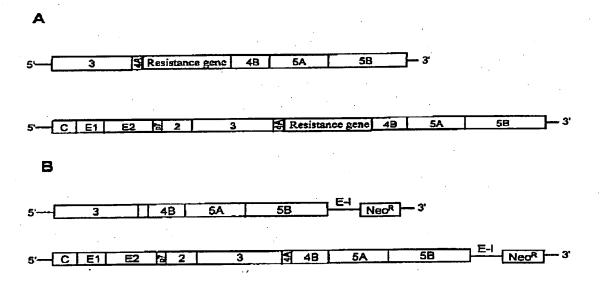


Fig. 5

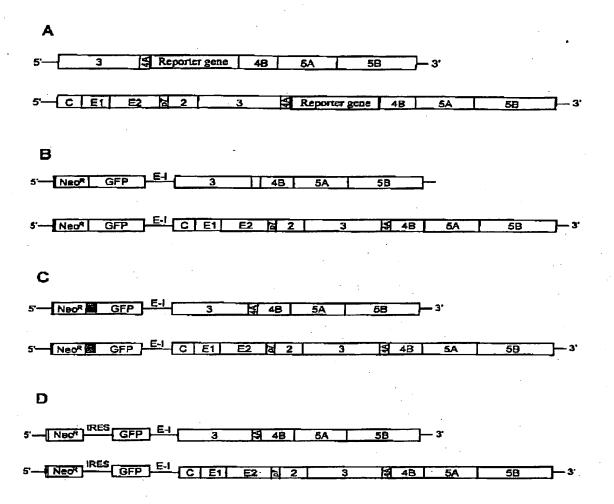


Fig. 6

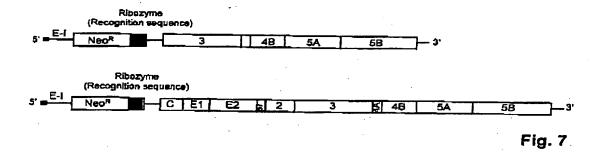




Fig. 8

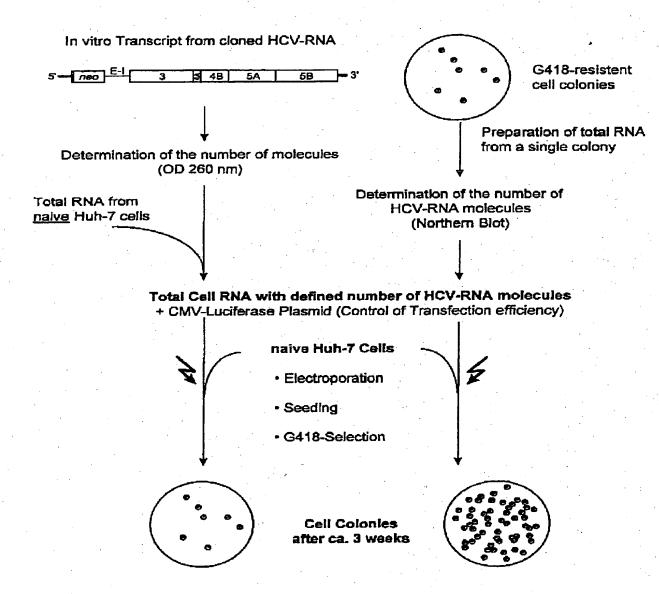
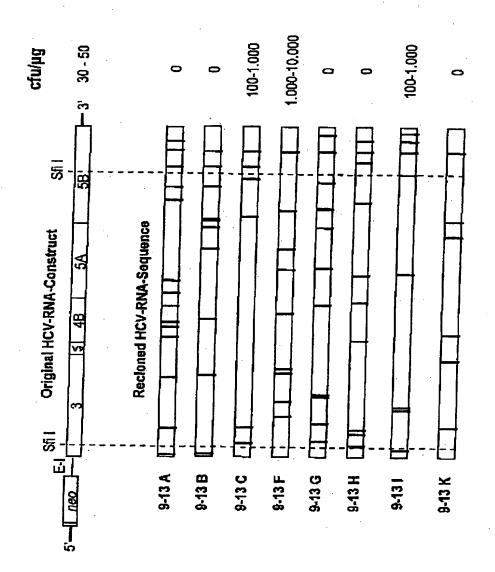
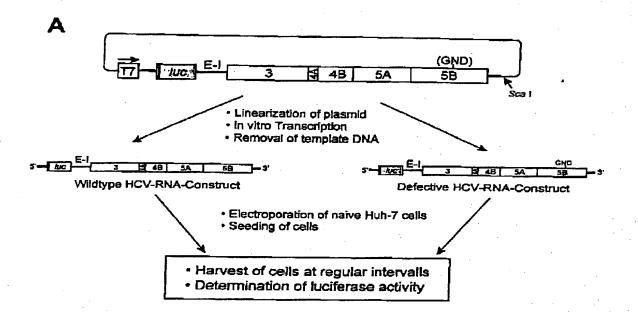


Fig. 9



Borden Ladner Gervais LLP | s.r.l.



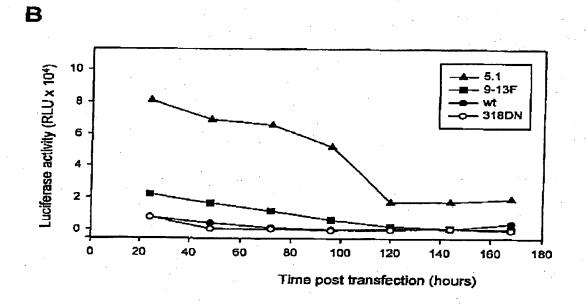
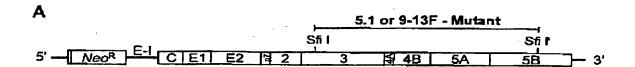


Fig. 11



В

G418-Concentration:	500 µg/ml	250 µg/ml	100 µg/ml
Number of Cell Colonies:	1	23	75

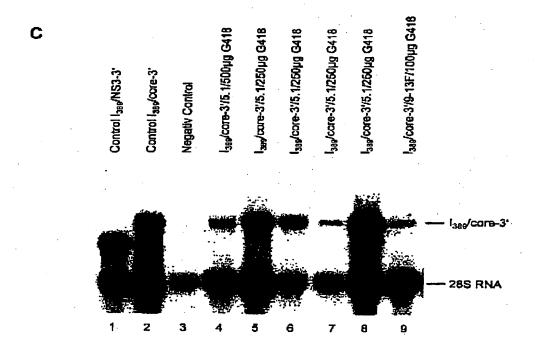
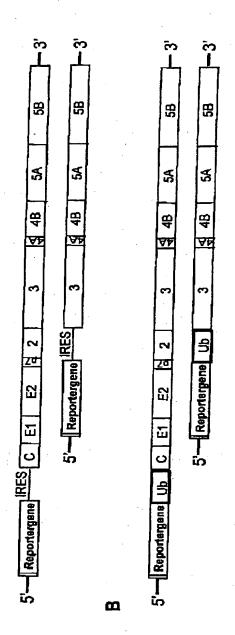
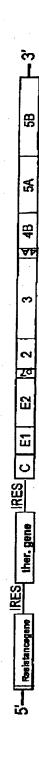
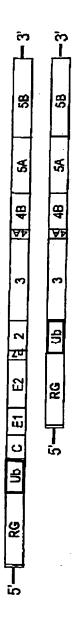


Fig. 12



Borden Ladner Gervais LLP / s.r.l.





## This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

□ BLACK BORDERS
 □ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
 □ FADED TEXT OR DRAWING
 □ BLURRED OR ILLEGIBLE TEXT OR DRAWING
 □ SKEWED/SLANTED IMAGES
 □ COLOR OR BLACK AND WHITE PHOTOGRAPHS
 □ GRAY SCALE DOCUMENTS
 □ LINES OR MARKS ON ORIGINAL DOCUMENT

## IMAGES ARE BEST AVAILABLE COPY.

☐ OTHER: \_\_\_

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY